AD_____

Award Number: W81XWH-06-1-0111

TITLE: Identification and Characterization of Prostate Cancer Associated Protein Biomarkers using High-throughput Mass Spectrometry

PRINCIPAL INVESTIGATOR: Gunjan Malik, Ph.D.

CONTRACTING ORGANIZATION: Cancer Therapy Research Center
San Antonio, Texas 78229

REPORT DATE: December 2007

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-12-2007	Final	15 NOV 2005 - 14 NOV 2007

4. TITLE AND SUBTITLE

Identification and Characterization of Prostate Cancer Associated Protein Biomarkers using High-throughput Mass Spectrometry

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0111

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Gunjan Malik, Ph.D.

E-Mail: gmalik@idd.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Cancer Therapy Research Center
San Antonio, Texas 78229

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Prostate cancer (PCa) remains to be the most common non-skin cancer in the US. Currently available screening tests for PCa including prostate specific antigen (PSA) test, digital rectal examination (DRE) and prostate biopsy, call for more accurate and non-invasive techniques to detect, diagnose, and stratify the disease based on molecular markers present in the body fluids. Using MALDI-TOF mass spectrometry protein "fingerprint" profiling, we generated decision tree algorithms to classify cancer from non-cancer. We have also devised strategies to isolate and identify protein biomarkers from the fingerprint profiles of PCa patients in the clinical gray-area where PSA fails to detect cancer. Identification of such cancer biomarkers will assist in development of better non-invasive diagnostic tools for prostate cancer and may also lead to better therapeutic targets.

15. SUBJECT TERMS
Stress, cerebral bloodflow, performance, vigilance, sustained attention, fatigue, vehicle driving, individual differences, workload

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	56	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

Page

Introduction…….. 4

Body…….. 4

Key Research Accomplishments…….. 6

Reportable Outcomes…….. 22

Conclusion…….. 27

References…….. 31

Appendices…….. 34

1. INTRODUCTION

Prostate cancer (PCa) remains to be the most common non-skin cancer in the US. Currently available screening tests for PCa including prostate specific antigen (PSA) test, digital rectal examination (DRE) and prostate biopsy, call for more accurate and non-invasive techniques to detect, diagnose, and stratify the disease based on molecular markers present in the body fluids. There has been an impressive emergence of mass spectrometry based technologies applied toward the study of such biomolecular markers of disease states. Our focus on utilization of such techniques towards prostate cancer will promise a better health and future for PCa patients. We have devised strategies to isolate and identify protein biomarkers from PCa patients in the clinical gray-area where PSA fails to detect cancer. Identification of such cancer biomarkers will assist in development of better non-invasive diagnostic tools for prostate cancer and may also lead to better therapeutic targets.

2. DESCRIPTION OF RESEARCH PROJECT:

Background:

Prostate Specific Antigen (PSA) testing has tremendously increased the detection of early-stage prostate cancer (PCa). However, a serum PSA value greater than 4.0 ng/mL warrants a biopsy that often indicates benign disease. On the other hand, recent assessments reveal an equally elevated risk (20-25% incidence) of PCa among men with serum PSA levels from 2.5 - 4.0 ng/mL. Our objective was to determine if serum protein-

expression profiles could be used to differentiate between benign and malignant prostate cancer in biopsy proven case (biopsy positive) and control (biopsy negative) patients with marginal clinical symptoms (serum PSA levels < 4.0 ng/mL).

Studies have demonstrated that high-throughput proteomic approaches for protein "fingerprint" profiling have tremendous potential for identifying biomarkers to improve prostate cancer diagnosis [Reviewed by (Petricoin et al, 2004; Semmes et al, 2006; Wright et al, 2005)]. A large number of proteins that are relevant in understanding the biological processes are expressed at low levels in the system. Therefore, there is a need for highly sensitive, high throughput methods to analyze a wide dynamic range of proteins. In order to improve the ability to "mine" the full depth of the proteome, we aimed to apply the UltraFlex™ MALDI-TOF/TOF instrumentation equipped with ClinProt robotic bead-based sample processing station (Bruker Daltonics). The MALDI-TOF instrument provides improved ability to mine deeper into the proteome, improved resolution/accuracy and the ability to achieve peptide/protein identification (Suckau et al. 2003).

The **SPECIFIC AIMS** of our original application were-

AIM 1. Serum cohort to identify prostate cancer (PCa) population with minimal clinical symptoms.

AIM 2. Discovery of protein biomarkers for the early detection of PCa in cohort.

AIM 3. Isolation and identification of the protein biomarkers.

AIM 4. Development of MS-assisted immunoassay for PCa diagnostics.

3. KEY RESEARCH ACCOMPLISHMENTS

3.A. Serum Cohort to Identify Prostate Cancer (PCa) Population with Minimal Clinical Symptoms.

Our studies were directed at the male population that present with marginal symptoms (such as low PSA levels and/or positive DRE) and who undergo biopsy. In our previous report, we had identified, collected and stored serum specimens from 185 patients with positive prostate biopsy. We had also collected a set of 223 serum samples from patients with negative biopsy. All the samples were stored at -80^{o}C in small aliquots ready to be used for this study.

For the first (pilot) phase of the project, 106 subjects were selected from the control group and 68 subjects were selected from the cancer group to form the sample cohorts for MALDI-TOF and MALDI-TOF/TOF analysis. No attempt was made to match the samples between cohorts for age, race, body mass index, or other risk factors. Attempt was made to maintain the date of biopsy no more than four weeks from the date of serum collection in this cohort.

Only pretreatment samples, obtained at the time of diagnosis of prostate cancer, were collected for use in this study. All samples were obtained from properly consented patients through the institutional review board approved protocols.

3B. Discovery of Protein Biomarkers for the Early Detection of PCa in Cohort.

In our original application, we proposed to establish the clinical utility of MALDI-TOF mass spectrometry approach to protein profiling and biomarker discovery, and employ it to the specific early detection objectives in PCa by applying these proven approaches to the characterization and sequence identification of promising biomarkers for detecting early cancer.

We targeted serum as a source of proteome in our studies because we have encouraging preliminary results that this source can serve as potential diagnostic assay in that it is routinely available clinically and demonstrates reasonable reproducibility in protein concentration.

However, the identification of biomarkers of cancer in a complex body fluid such as blood and/or serum requires effective sample preparation prior to mass spectrometry-based analyses (Semmes et al. 2006). An effective sample preparation technique would not only significantly reduce the complexity of the samples, but also eliminate the abundant proteins such as albumin, immunoglobulins etc. from the samples, hence concentrating the low molecular weight and low abundance proteins, enhancing their eventual visualization on the MS platform.

The objectives of this aim were achieved by the following sub-aims-

3.B.1. Development and Evaluation of Methodologies for Protein Profiling.

As indicated in our previous report, a significant portion of our efforts were spent in the first year to optimize the strategies to enrich low abundant proteins to allow for their identification by mass spectrometry. For the most part, we utilized pooled human serum set referred to as QC (quality control serum) as our evaluation sample set to develop methodologies for up-front fractionation of serum for MALDI profiling. After careful assessment of two parameters- whether serum depletion prior to fractionation improves the detection of the proteins, and which fractionation strategy provides the most differential capture of these proteins between case and control, we established protocol(s) for paramagnetic bead-based fractionation for automation of the techniques on the ClinProt robotic workstation. The employment of functionalized magnetic bead-based techniques in conjunction with mass spectrometry allows for much shorter sample processing times and automatic workflows for efficient reproducibility. Subsequent high-resolution MALDI-TOF allows for highly sensitive analyses of the detected proteins and/or peptides.

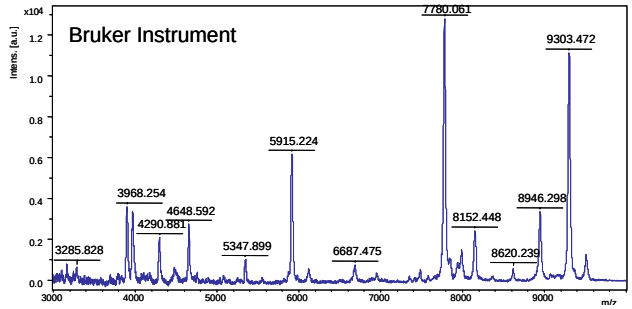


Figure 1. A typical spectrum of quality control serum (QC) on the Bruker UltraFlex platform displays the presence of three major serum peaks obtained on IMAC beads that were used for protocol optimization.

We designed our approaches to target two different profiling strategies- "Top-down" and "bottom-up" approach. In the *top-down* proteomics approach, we aimed to resolve the whole proteins from un-fractionated serum samples utilizing paramagnetic

beads prior to mass spectrometry analysis. In the *bottom-up* proteomics approach, complex protein mixtures can be enzymatically digested prior to separation and differential expression determined using mass spectrometry.

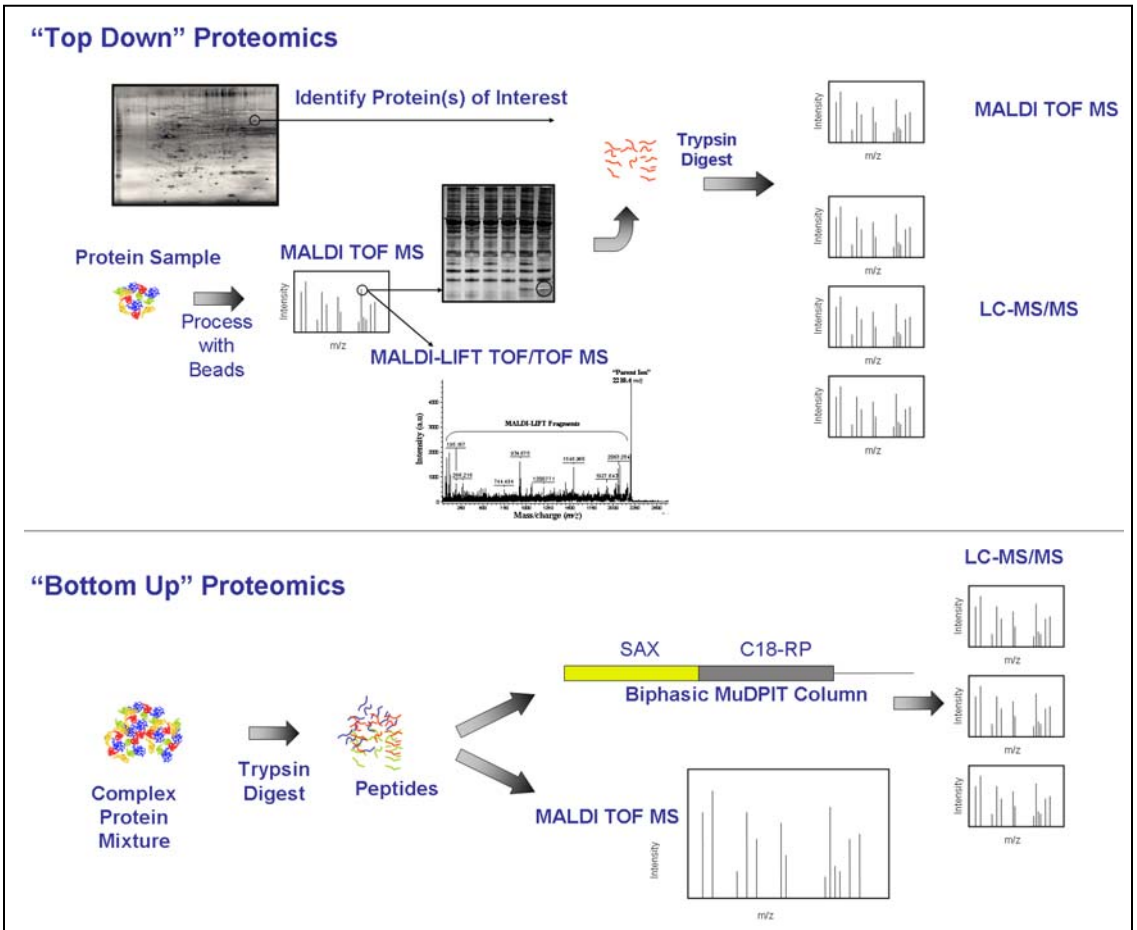


Figure 2. Two major approaches to clinical proteomics. **A**: In the Top-down proteomics approach, whole proteins are pre-fractionated via various gel and non-gel based techniques. Intact proteins of interest are then subjected to subsequent MS-based analyses using either single or tandem mass spectrometry. **B**: The bottom-up proteomics approach utilizes primarily non-gel based fractionation of peptides generated from complex protein mixtures. Digest-generated peptides are then subjected to mass spectrometry analysis. The approach can involve simple peptide mass profiling as well as quantitative tandem mass spectrometry to yield protein identification and relative protein concentration. *Modified from* (Semmes and Malik et al, 2006).

In all of these conceptual approaches (see **Figure 2**), front-end sample fractionation and separation strategies are required to reduce the complexity of native clinical sample such as serum. After careful assessment of both the approaches in the first year and using the quality control serum (QC) for a stepwise evaluation process based

upon achieving specific parameter objectives that measure the sensitivity, mass accuracy, signal-to-noise, resolution and reproducibility of the instrumental process, we first focused on the protein expression profiling using the top-down approach (**Figure 3**, *In braces*), as described in the following section.

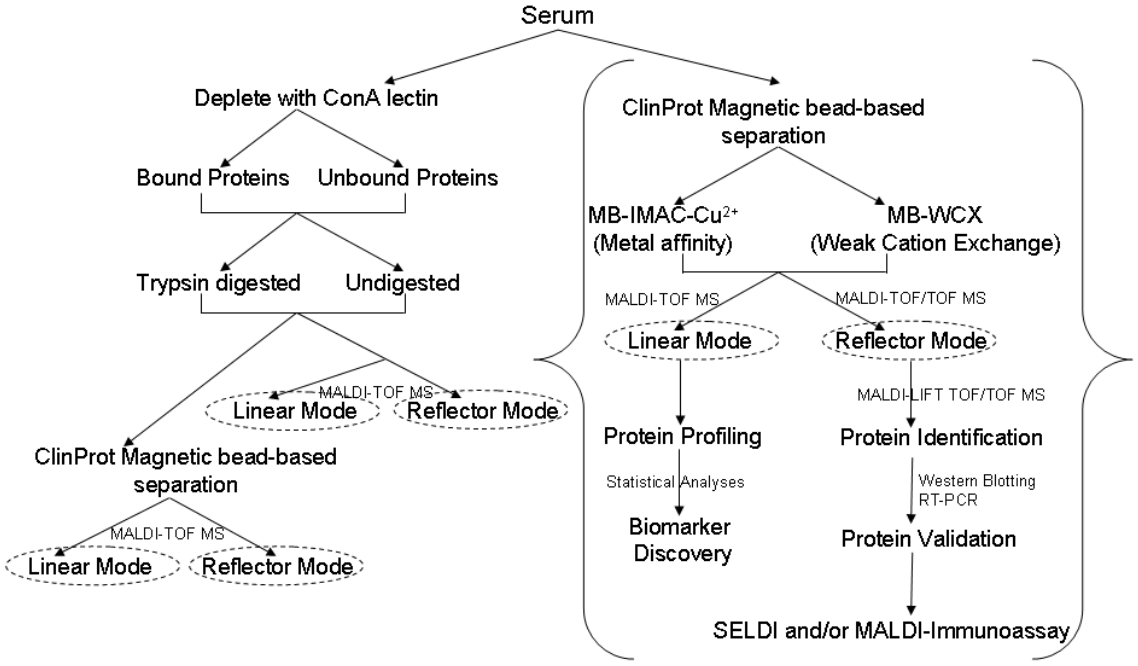


Figure 3. Experimental flow chart for protein profiling using mass spectrometry.

3.B.2. "Top-Down" Approach for Protein Profiling.

Protein expression profiling using MALDI-TOF approach has seen a wide application to many disease sites including prostate cancer. Our laboratory and others have been employing a combination of chromatographic paramagnetic beads and MALDI TOF/TOF MS to present a powerful and sensitive analysis of pre-fractionated samples (reviewed by [Pusch and Kostrzewa, [2005]]). The paramagnetic beads allow for reasonable high throughput processing and reproducible fractionation of

proteins/peptides, followed by MALDI-TOF MS analysis (**Figure 3**). Since the introduction of this technology to the field, the technique has been widely used for single or multidimensional separation of proteins/peptides on the beads. The fractions are then spotted on target plates for MALDI-TOF analysis [Villanueva et al., [2004]].

We utilized metal-binding (IMAC-Cu) or cation-binding (WCX) paramagnetic beads for this approach. Based on our past experience with the magnetic bead-based separations, we had chosen these two beads to be most functional in capturing a large number of protein/peptide "peaks" while establishing a case versus control differential prior to mass spectrometry analysis.

Details of the protocol are given in the attached manuscript. Briefly, whole un-fractionated serum was incubated with the magnetic beads and non-bound components removed by subsequent wash steps. Sample tubes containing the magnetic beads were mixed and rinsed robotically on a magnetized surface on a Bruker ClinProt robotis workstation. All sample processing (binding, washing, elution, matrix addition, plate spotting) were performed with this automated system. Eluted proteins/peptides from the beads were then spotted on an AnchorChip sample target platform (384 spots). Profile spectra were acquired in the linear mode as well as reflector mode on an Ultraflex™ MALDI-TOF/TOF instrument (**Figure 3**). The performance of this system and any optimization were based upon the same parameters optimized for the display of key QC peaks (**Figure 1**). Suitable protein/peptide peaks were then analyzed by MALDI-LIFT

TOF/TOF MS to identify the corresponding proteins by database search. *Details of the research design and methods are given in the attached manuscript*.

The system was evaluated using defined case (n = 68) and control (n = 106) samples from our serum cohort (as described in Section 3A) to examine the ability of the UltraFlex™ to achieve correct classification using algorithms that are available in the Bruker MALDI-TOF software suite as well as newer classification approaches. The analysis of the MALDI data in house was accomplished by Dr. Malik. In parallel we also sent the data to Dr. John Cornell (UTHSCSA) for analysis using Decision Tree algorithm and other classification approaches (*see letter of collaboration in appendix*).

Initial analysis of MALDI data *processed* by the Bruker software for baseline subtraction, peak alignment, peak selection, normalization of intensity and mass/charge calibration, using a Decision Tree algorithm yielded **73.5% sensitivity** and **93.8 % specificity** for classifying cancer and non-cancer cases with an area under the ROC curve of 0.94 (*see attached manuscript for details*). Analysis of *Raw* or *Un-processed* MALDI TOF data is under progress and will be included in the manuscript.

The proposed system for triage allowed for a rapid decision regarding the pursuit of protein profiling and/or individual biomarker discovery approaches. Subsequent analysis with a larger well-designed clinical sample set will test the utilization of this approach in cancer diagnostics (see section *3.B.4.*).

3.B.3. "Bottom-up" Proteomics Approach

As described in detail in our previous report, we aim to apply the bottom-up proteomic approach towards the capture the glycoproteins and/or peptides. The captured glycoproteins will then be digested with trypsin or similar enzyme to generate peptide fragments from each protein (See **Figures 2 and 3**). Proteomic or "peptidomic" profiling of these digest generated peptides will result in fast and efficient profiling of high molecular weight proteins and therefore, overcome the limitations on the resolvable mass range in MALDI mass spectrometry. This will also lead to fast and highly sensitive protein identification from the protein fragments in the MALDI-LIFT TOF/TOF mode. This approach will integrate high-resolution separation of digest-generated peptides with increasingly sophisticated mass spectrometry for bottom-up differential identification.

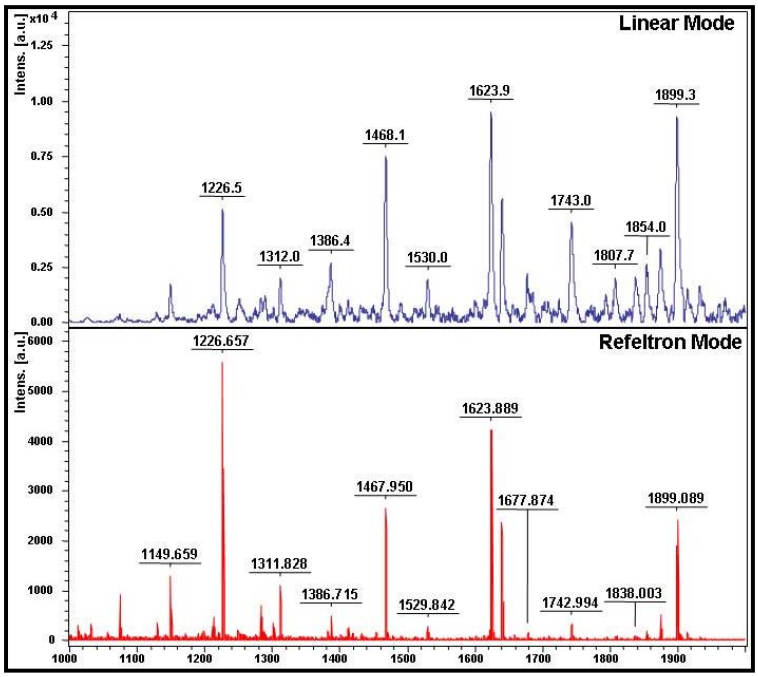


Figure 4. Pooled quality control serum was digested with Trypsin and analyzed by MALDI-TOF MS after purification with ClinProt MB-ConA beads. The sample fractions were run on both Linear (upper) and Reflector (lower) mode.

Initial evaluation of this approach using QC serum has been achieved as reported last year (**Figure 4**). We will now utilize samples from the *same* serum cohort for case (n

= 68) and control (n = 106) comparison on this platform. In an effort to maintain the number of freeze/thaws the same for both the studies described in section *3.B.2* and in this section, duplicate aliquots of this sample set were prepared and stored at -80°C for these studies.

3.B.4. Validation of Protein Profiling in Larger Sample Cohort.

The initial proof-of-concept protein profiling study as decsribed in Section *3.B.2.* directs us to test the MALDI-TOF platform in a much more sofisticated and larger case versus control study (n ≈ 300 each). A critical component of our approach would be the overall study design. Acquiring serum samples just before biopsy would assure that each sample would be handled in a similar manner. Variables related to serum clotting and storage times before freezing would be minimized. Each sample will be appropriately aliquoted at this initial step to minimize freeze-thaw cycles. The need for a "normal" population aquired from healthy volunteers, which can be particularly difficult to match with samples acquired in the clinic, is not necessary for this study. Instead, pooled reference serum samples (QC) would be included within the analytical process to ensure the reproducibility of the process and the quality of the spectra generated. The bead capture steps and spotting of the samples for MALDI analysis are all fully automated. Since the sample size would be much larger in this cohort, throughoput of the MALDI platform would be maintained by incorporating a Twister attachment (Bruker Daltonics) a robotic arm that automates the reading of the Anchor chip for an overall throughput in

multiples of 384 samples, thus automating the sample processing pass-through from ClinProt robot to UltraFlex™.

We will utilize the samples collected and maintained by Dr. Ian M. Thompson (see letter in appendix). His laboratory would be our source of samples for the second phase of validation and a letter of collaboration is included for the same.

3C. Identification of Diagnostic Proteins/Peptides Using the UltraFlex™ TOF/TOF.

The recent introduction of MALDI LIFT-TOF/TOF mass spectrometry through Bruker Daltonic's release of the UltraFlex™ system has provided the proteomics community with unprecedented capabilities [for review see (Suckau et al. 2003)]. When operated in the TOF/TOF mode the UltraFlex™ achieves very high resolution, accuracy and signal to noise and effective tandem mass spectrometry for protein identification (**Figure 5**).

After initial analysis in the linear mode, target proteins/peptides of interest are visualized for further identification and characterization. When the protein peak has been targeted, identification is achieved with the UltraFlex™ which employs ion potential lift (LIFT) technology in a MALDI-TOF/TOF platform for highly sensitive (attomolar range) and accurate tandem mass spectrometry for peptide mass fingerprints (PMFs) (**Figure 5**).

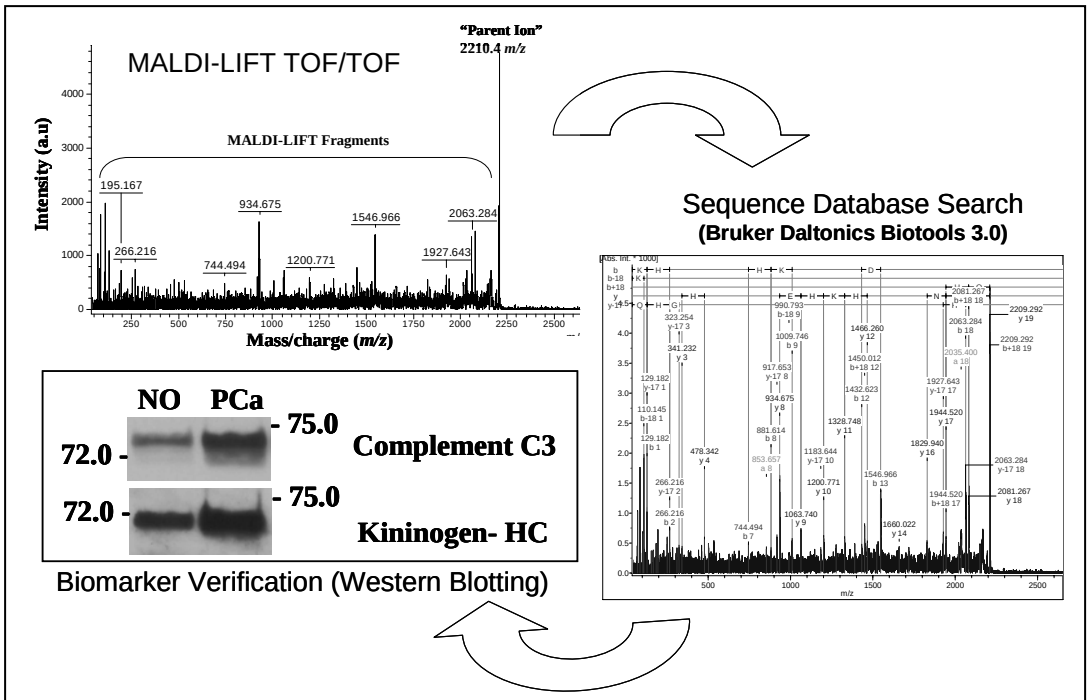


Figure 5. Experimental flow-chart for protein identification using MALDI-LIFT-TOF/TOF approach. Protein peak of interest is "gated" for MALDI-LIFT and identified by sequence search using Biotools 3.0. Subsequent verification of the protein identity is performed by Western blotting of MALDI eluates.

In a proof-of-concept study, we applied MALDI-LIFT TOF/TOF approach to some of the major proteins/peptides that displayed a significant differential expression between cases vs. controls ($p < 0.05$) in the study described in section *3.B.2.* (See attached manuscript for details). In brief, to identify some of the protein peaks ($p < 0.05$) generated by the MALDI-TOF, MALDI-LIFT-TOF/TOF spectrometry was applied to the eluates from the magnetic beads in the presence of CHCA matrix. Fragment ion spectra were first analyzed with FlexAnalysis 3.0 (Bruker Daltonik GmbH, Germany). Peptide mass fingerprints generated by the MALDI-LIFT approach were used for MASCOT (Matrix Science, London, UK) search employing Biotools 2.2 (Bruker Daltonics). Initial analysis revealed the identity of some of the key proteins overexpressed in case vs. the control set in more than one peptide fragments. The identified proteins were verified by Western blotting of the MALDI eluates with antibodies specific to the proteins. *The*

details of results and methodologies used in this section are given in the attached manuscript.

A major advantage to the combined ClinProt UltraFlex™ system is the ability to directly scale up for isolation and purification prior to applying sequence identification efforts. Specifically, the identification of small mass (<10,000 Da) proteins/peptides can be achieved; a mass range not effectively mined by traditional gel based approaches. For large sized and/or hard to fragment proteins, protein fractions generated by the MALDI beads will be trypsinized and applied either to combined Laser Induced Dissociation (LID) or LC-MS/MS for protein identification. We have had good success with the protein identification using a combination of SELDI, affinity chromatography, gel filtration chromatography and tandem mass spectroscopy (Malik et al, 2007; Malik et al, 2005). Proteins/peptides that are very large are not likely to be analyzed with MALDI-LIFT approach. Thus, the combination of the two technical approaches gives us much improved "coverage" with respect to the range of proteins that can be identified.

The profiling data generated by MALDI-TOF MS will also be used to identify paired samples that greatly over-express or under-express the targeted biomarker(s) of interest. The selected paired samples will then each be subjected to further verification by Western analysis.

The identification of protein/peptide components of a fingerprint protein profile is critical to both validating the utility of the disease fingerprint pattern and to maximizing

the utility of the individual biomarker. For instance knowing the composition of the proteins/peptides that are over-expressed and/or under-expressed that comprise the diagnostic pattern will provide a surrogate marker for the profiling assay and enable optimization of the "diagnostic" platform. In addition identification of all the proteins/peptides that comprise the diagnostic pattern will provide the basis for development of a multiplexed immuno-assay; which could potentially enter into the clinic as a diagnostic test more rapidly (See next section for details).

3.D. Development of MS-assisted Immunoassays for PCa Diagnostics.

We have previously identified biomarker protein/peptides that comprise disease-specific signature profiles. Each of these biomarker proteins are then verified and validated using antibody-based assays. Two such biomarkers, Apolipoprotein-AII and Histone H2B have reached the point of designing pre-validation studies (Malik et al, 2007; Malik et al, 2005).

Upon successful fulfillment of Aims 2 & 3 (Section 3B and 3C) we will acquire or develop antibodies to all the identified proteins and a dilution end-point Western analysis will be conducted for each (**Figure 6**) to determine

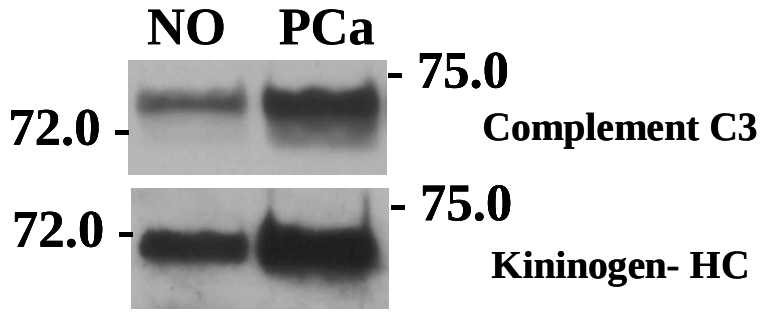


Figure 6. Western Analysis of paired case (PCa) and control (NO) samples generated by WCX-magnetic beads on ClinProt robot.

the relative difference in expression levels between each of the paired case/control samples.

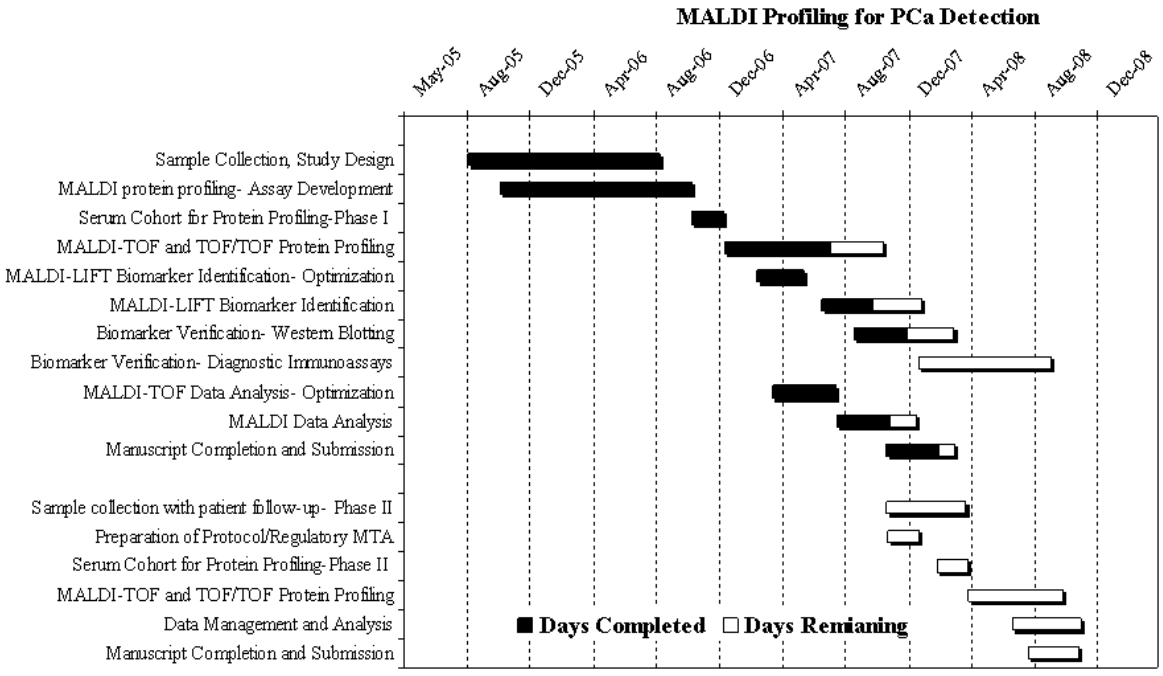
Most of the identified proteins were much larger than the protein/peptide fragment peaks identified on MALDI-LIFT. However, most of the proteins were identified from as many as four unique peptides, thus increasing the overall percent coverage of the identified protein peak. To verify that the identified biomarker is indeed the same peak or set of peaks, as seen in the MALDI trace, the antibodies will be used to immuno-deplete the sample prior to analysis by MALDI. This trace will be compared to a spectral trace of the same sample immuno-depleted with a non-specific protein of the same isotype. In this comparison the immuno-depletion with the specific antibody should show a decrease in the diagnostic MALDI protein peak or protein fragments. Since we have verified that antibodies specific to the identified proteins were effective at recognizing the respective protein on Western blots, we are now in the process of using these reagents to selectively immuno-deplete the proteins and their respective MALDI peaks from the sera.

The next verification study utilizes the SELDI or MALDI-based immuno-assay capabilities. The ability of SELDI to function as an accurate immuno-assay tool has been demonstrated (Malik et al, 2007; Malik et al, 2005; Xiao et al, 2001; Xiao et al, 2000). In short, the specific antibodies are attached to a derivitized ProteinChip® surface and the treated ProteinChip® is then reacted with sample. Antibody specific antigens are then detected as peaks corresponding to antigen mass. This approach has the added advantage

of being able to examine antigen fragments apart from whole protein. Protein G or Protein A coated magnetic beads can also be cross-linked to specific antibodies for high affinity, quantitative capture of selective biomarkers on the MALDI platform. Positive results in these two steps will warrant a mini-validation using the SELDI or MALDI immuno-assay. Specifically, 50 cases and 50 controls will be processed using the antibody-specific proteinchip or magnetic beads and examined for expression levels that correspond to the original MALDI and also provide promising discriminating power.

The ability of the identified and validated biomarkers to diagnose prostate cancer, especially in sample groups where PSA fails to detect cancer (clinical gray area), would be tested using large sample sets on MALDI and SELDI-based immunoassays using sample cohorts from Dr. Ian Thompson. *Data generated from the immuno-depletion and mass spectrometry-assisted immunoassays will be used to report it in a manuscript form.*

Timetable for the Proposed Studies



Work Planned for Remaining Year

1. We will finish the MALDI-TOF and TOF/TOF Protein Profiling in serum cohort designed for the proof-of-concept study using the "bottom-up" peptidomics approach. Serum aliquots with the *same* freeze-thaws as in the study reported here are prepared and stored at -80°C. *A manuscript describing these studies is in preparation* (see attached appendix).
2. We will finish MALDI-LIFT biomarker identification, especially in the high molecular weight range (> 10,000Da) using combinatorial approaches. All the identified biomarkers will be verified by Western blotting.
3. Biomarker Verification- to develop diagnostic immunoassays, antibodies to identified biomarkers will be tested on SELDI and/or MALDI-based immunoassays and then validated by larger sample sets. This work will be performed in collaboration with Dr. Ian Thompson.
4. MALDI Data Analysis is being conducted in collaboration with Dr. John Cornell, UTHSCSA (see attached letter). Data generated from the *raw* un-processed MALDI data of the pilot experiment will be included in the attached manuscript.
5. Sample collection with patient follow-up- Phase II- This work will be done in collaboration with Dr. Ian Thompson. His laboratory is developing a prospective collection of samples from patient follow-ups which will be used for second phase of Aim2 in our study.
6. Sensitivity and specificity of the generated algorithm(s) would be tested and validated using this independent test set of the prospectively collected samples. *This phase will provide the data needed for submission in a manuscript form.*

4. REPORTABLE OUTCOMES

4.1. As reported last year, serum samples from 223 patients with negative prostate biopsy (controls) and 185 patients with positive prostate biopsy (cases) in the clinical gray area of diagnosis (PSA < 4.0 ng/mL; Abnormal DRE and/or elevated PSA etc.) were collected and stored in our serum repository for use in this study.

4.2. Initial evaluation of the pre-fractionation of serum samples prior to MALDI-TOF MS was performed. Based on the results of the initial pilot-experiments, bead-based capture of whole un-fractionated serum prior to MS analysis generated the best outcome.

4.3. A proof-of-concept MALDI-TOF profiling study was performed in carefully selected case (n = 68) versus controls (n = 106) using IMAC-Cu and WCX paramagnetic beads and processed in both Linear and Reflector mode on the Bruker Ultraflex platform.

4.3.1. All the samples were processed in a randomized format in duplicates using robotic magnetic bead based enrichment with MB-IMAC-Cu and MB-WCX beads and analyzed on Bruker Ultraflex III MALDI-TOF mass spectrometer in both linear and reflector mode.

4.3.2. About 1500 spectra (348 spectra generated from each data set) were analyzed for peak intensity normalization, baseline subtraction, calibration and peak picking using Flex Analysis 3.0 software.

4.3.3. ClinProt 2.0 software, used for initial analysis of sensitivity and specificity of the protein peaks generated with the software generated poor to modest classification.

4.3.4. Application of Adaboost algorithm with a J48 decision tree algorithm with pruning to the MALDI *processed* data generated the strongest combination of **73.5% sensitivity** and **93.8% specificity** for classifying cancer and non-cancer cases with an area under the ROC curve of 0.94. This was conducted in collaboration with Dr. John Cornell (a letter is attached). Results of the analysis of the *processed* data (completed) and *raw* data (In progress) will be incorporated in the attached manuscript.

4.4. For protein identification, MALDI-TOF protein peaks with the best differential in case vs. control ($p < 0.05$) were analyzed my MALDI-LIFT in the TOF/TOF mode. Initial results identified three major proteins from the WCX eluates- Complement component C3, Fibrinogen-alpha and Kininogen, in more than one peptide "peaks" (total protein coverage ≈ 30%), overexpressed in cases as compared to controls. None of the differential protein peaks from IMAC eluates could be identified probably due to posttranslational modifications. Efforts are underway to devise strategies to identify and verify all the differential peaks identified from MALDI platform.

4.5. Western analysis was performed in paired overexpressing and underexpressing case and control sample eluates from MALDI beads to verify the protein identities.

4.6. During my postdoctoral training at Eastern Virginia Medical School (EVMS), I was offered a position at the newly developed Division of Molecular Pathology at the Cancer Therapy and Research Center's (CTRC) Institute for Drug Development (IDD) in San Antonio, Texas. I joined there as a Senior Research Associate from Oct 2, 2006.

4.7. Department of Defense was requested to transfer the postdoctoral traineeship award from EVMS to CTRC which was successfully completed in March, 2007.

4.8. Since several months were spent in my re-location, award transfer and setting up the new laboratory at the Molecular Pathology Division, CTRC, a no-cost 12 month extension was separately requested for the grant period.

Peer-reviewed publications related to proteomics:

1. Dale McLerran…**Gunjan Malik**, EPSIC members and O. John Semmes. SELDI-TOF-MS whole serum proteomic profiling with IMAC surface does not reliably detect prostate cancer. *Clin. Chem.* 2007 Nov 16. *In Press*.

2. Dale McLerran…**Gunjan Malik**, EPSIC members and O. John Semmes. Analytical validation of protein expression profiling for diagnosis of PCa; Sources of sample bias. *Clin. Chem.* 2007 Nov 2. *In Press*.

3. **Gunjan Malik**, Elizabeth Rojahn, Michael D. Ward, Mathew B Gretzer, Alan W. Partin, O. John Semmes, Robert W. Veltri. SELDI Protein Profiling of Dunning R3327 Derived Cell Lines: Identification of Molecular Markers of Prostate Cancer Progression. *The Prostate*. 2007 Aug 17; 67(14):1565-1575.

4. **Malik G**, Ward MD, Gupta SK, Trosset MW, Grizzle WE, Adam BL, Diaz JI, Semmes OJ. Serum Levels of an Isoform of Apolipoprotein A-II as a Potential Marker for Prostate Cancer. *Clin. Cancer Res.* 2005 Feb 1; 11(3):1073-1085.

Published abstracts related to proteomics:

5. **Gunjan Malik**, Saurabh K. Gupta, Michael D. Ward…O. John Semmes and Jose I. Diaz. Proteomic Analysis of T24T Derived Bladder Cancer Cell Lines Using Differential In-Gel Electrophoresis and LC-MS/MS. Annual AACR Meeting (April 14-18, 2007, Los Angeles, CA).
6. Rojahn Elizabeth, Sumit Isharwal, **Gunjan Malik**, Alan W. Partin, Robert W. Veltri. A novel membrane p17 protein biomarker is overexpressed in metastatic human and rat (Dunning) prostate cancer cell lines, human prostate tissues and serum. Annual AACR Meeting (April 14-18, 2007, Los Angeles, CA).
7. Robert W. Veltri, **Gunjan Malik**, Elizabeth Rojahn, Cameron Marlow, Michael Ward, Alan W. Partin. "PBOV1 (UC28): Molecular characterization and assessment as a serum marker for detection of prostate cancer (PCa)". Annual AACR Meeting (April 16-20, 2005, Anaheim, CA)

Non peer-reviewed publications related to proteomics:

1. Grizzle, WE, Semmes, OJ, Bigbee, WL, **Malik, G**, Miller, E, Manne, B, Oelschalger, DK, Zhu, L, Manne, U. Use of high throughput mass spectrographic methods to identify disease processes with emphasis on SELDI-TOF-MS methods. *In:* George Patrinos, Wilhelm Ansorge (ed.), Molecular Diagnostics, Elsevier Press., June 06, 2005 Chapter 17: 211-222. ISBN: 0-12-546661-7
2. Lisa H. Cazares, Richard R. Drake, **Gunjan Malik** and O. John Semmes. SELDI-TOF profiling for clinical diagnostic assay development. *In*: Fotini T. Stathopoulou (Editor), Genome and Proteome in Oncology, Nova Sciences Publishers, Inc., Mar 30, 2005 Chapter 6: pp. 113-127. ISBN: 1-59454-285-6.

Peer-reviewed publications related to the original grant application:

1. Drake RR, Schwegler EE, **Malik G**, Diaz JI, Block T, Mehta A, Semmes OJ. Lectin capture strategies combined with mass spectrometry for the discovery of serum

glycoprotein biomarkers. *Mol. Cell. Proteomics*. 2006 Oct; 5(10):1957-67.

2. O. John Semmes, **Gunjan Malik** and Mike Ward. Application of Mass Spectrometry to the Discovery of Biomarkers for Detection of Prostate Cancer. *Journal of Cellular Biochemistry*. 2006 Jun 1; 98(3):496-503. Review.

Published and/or submitted abstracts related to the original grant application:

1. **Gunjan Malik**, Lisa H. Cazares, O. John Semmes and Jose I. Diaz. Identification and Characterization of Prostate Cancer Associated Protein Biomarkers using High-throughput Mass Spectrometry. Annual AACR Meeting (April 12-16, 2008, San Diego, CA).
2. **Gunjan Malik**, Lisa H. Cazares, Kali Makedou, Saurabh K. Gupta, Shamina G. Mitchell, Mary Ann Clements, Tarek O. Kandil, Brian P. Main, Richard R. Drake, O. John Semmes and Jose I. Diaz. Identification and Characterization of Prostate Cancer Associated Protein Biomarkers using High-throughput Mass Spectrometry. Department of Defense PCRP Innovative Minds in Prostate Cancer Today (IMPaCT) Meeting (September 5-8, 2007, Atlanta, Georgia).

Manuscripts *in progress*:

1. **Gunjan Malik**, Lisa H. Cazares, Saurabh K. Gupta, John. E. Cornell, Kali Makedou…O. John Semmes and Jose I. Diaz. Identification of prostate cancer associated protein biomarkers in the clinical "grey area" using high-throughput mass spectrometry. *In Progress (Manuscript Appended).*
2. **Gunjan Malik**, Saurabh K. Gupta, and Jose I. Diaz. Identification of bladder cancer markers in the T24 model system by 2D-DIGE and LC-MS/MS analysis. *In Progress*
3. **Gunjan Malik**, Saurabh K. Gupta, Michael D. Ward…O. John Semmes and Jose I. Diaz. Proteomic analysis of T24T derived bladder cancer cell lines correlates enhanced BMP signaling to liver metastasis. *In Progress*

4. Saurabh K Gupta, **Gunjan Malik**, James F. Courage and Jose I. Diaz. Characterization of gene expression signatures during prostate cell differentiation in normal prostate cell epithelium. *In Progress*

Grant application(s) applied for/under review related to proteomics:

1. PA-06-299- National Cancer Institute (NCI), Exploratory Studies in Cancer Detection, Diagnosis, and Prognosis (R21). Title- "*A comprehensive genomic and proteomic analysis of molecular markers contributing to the metastatic ability of cancer cells and their subsequent clinical validation*". Earliest Anticipated Award Date- June, 2008.
2. 2008 San Antonio Area Foundation Research Grant. Title- "*Development of Molecular Markers of Cancer Metastasis*". Earliest Anticipated Award Date- May, 2008.

5. SUMMARY/CONCLUSIONS

Prostate cancer (PCa) is the most frequently diagnosed cancer in men. With an estimated 27,050 deaths from PCa, it is a leading cause of cancer death in men. The "gold standard" diagnostic marker for PCa is prostate specific antigen (PSA) and the rapid incorporation of aggressive PSA testing has resulted in a dramatic reduction in the identification of advanced stages of PCa as well as deaths secondary to PCa (McDavid et al, 2004; Carter et al, 2004). However, increasing number of reports are emphasizing the limitations of the maker in prostate cancer diagnosis. More than 90% of all PCa are discovered in the local and regional stages with their 5-year survival rate reaching almost 100%. However, the survival rate drops to 33% when PCa has spread to distant sites. Approximately 40,000 men die each year with PCa metastasis (Jemal et al, 2007).

Recent findings suggest that 15-25 percent of men with a "normal" PSA level of <4.0 ng/mL have had prostate cancer, which therefore underscores the need to consider fundamental changes in the approach to diagnosing prostate cancer (Thompson et al, 2004). Several other calculated parameters, such as PSA density, PSA transition zone density, PSA velocity or age- and race-specific PSA ranges, were only partially successful in enhancing the specificity of PSA (Catalona et al, 2000; Nixon, 1997; Thompson et al, 2006). Expression profiling and proteomics have the potential to transform the management of prostate cancer, identifying new markers for screening, diagnosis, prognosis, monitoring and targets for therapy (Masters, 2007).

In this grant application, we proposed to design and conduct carefully planned protein profiling studies for the discovery of new and novel biomarkers in serum of patients with a “normal” PSA, which could be used to differentiate between biopsy-proven cases and controls. Serum was collected and stored from patients who presented marginal clinical symptoms (PSA < 4.0 ng/mL and/or abnormal DRE etc.) and a subset of 68 cases and 106 controls were subjected to MALDI-TOF and MALDI-TOF/TOF mass spectrometry protein profiling using two different types of paramagnetic bead-based separation techniques. Samples were run in both linear and reflector mode on the Bruker Ultraflex platform. Identified “peaks” were utilized to develop classification algorithms using both in-house as well as other classification approaches (in collaboration with Dr. John Cornell, University of Texas Health Science Center, San Antonio).

Protein/peptide peaks displaying significant differential expression ($p < 0.05$) between cases vs. controls were also subjected to MALDI-LIFT TOF/TOF for protein identification. Identified proteins were then verified in paired over-expressing and under-expressing case vs. control samples using Western blotting with antibodies specific to the identified proteins. These antibodies will now be used to develop SELDI and/or MALDI-based immunoassays. Identification of all the proteins/peptides that comprise the diagnostic pattern will provide the basis for development of a multiplexed immuno-assay; which could potentially enter into the clinic as a prognostic test more rapidly.

Further analysis of larger, well designed case vs. control sample sets, stratified by disease stage and grade, in patients with “normal” PSA is under progress in collaboration with Dr. Ian Thompson at the University of Texas Health Science Center, San Antonio TX. This sample set will be used to *challenge* the algorithms developed by the initial studies as well assess the robustness of the platform. The samples will also be used to validate the mass spectrometry-assisted diagnostic immunoassays.

The incorporation of expression differences of serum proteins into a diagnostic platform may prove to be an important parameter in the realization of challenging objectives of prostate cancer diagnostics. The identification of the individual differentially expressed proteins that comprise the diagnostic expression profile is essential to facilitating real progress in the development of a robust accurate diagnostic platform, because classic measurements of serum levels of proteins that comprise the profiles will help to stabilize/normalize the profile from patient to patient. In addition, if

the proteins are identified and specific high affinity antibodies are generated to them, then more direct and potentially less expensive immunodiagnostic methods for analysis can be developed. Identified marker, or marker panel, may not replace the need for PSA screening and prostate biopsies, but would improve their use and help to minimize unnecessary biopsies.

We report that a small sample set (n = 68) of patients with PCa could be distinguished from benign disease and healthy men (n = 106) with a **73.5% sensitivity** and **93.8 % specificity**. The resulting false-negative rate of this algorithm may not supplant the existing capabilities of PSA "cut-off" value. However, it is notable that these "fingerprint" profiles retain the discrimination between disease and non-disease when PSA levels are < 4.0 ng/mL or in other words, in cases of PCa in which PSA would have failed to detect the disease. Thus, the use of a robust fingerprint pattern, in combination with PSA may extend the utility of this test. This is especially important in light of the recent results from the prostate cancer prevention trial (PCPT) showing that a significant number of advanced cancers go undetected in patients with "normal" PSA values.

In addition, identification of the proteins/peptides that comprise of these diagnostic fingerprints would result in the incorporation of an immune-based assay for the identified and verified proteins for the development of a more robust assay platform than mass spectrometry alone. The immunodiagnostic assays can then be utilized to detect the actual levels of these markers in prostate cancer patient sera having PSA < 4.0 ng/mL, thus extending the utility of current blood testing for PCa.

5. REFERENCES CITED

1. Carter HB. Prostate cancers in men with low PSA levels--must we find them? *N Engl J Med* 2004; 350:2292-4.
2. Catalona WJ, Southwick PC, Slawin KM, Partin AW, Brawer MK, Flanigan RC, Patel A, Richie JP, Walsh PC, Scardino PT, Lange PH, Gasior GH, Loveland KG, Bray KR. Comparison of percent free , density, and age-specific cutoffs for prostate cancer detection and staging. *Urology* 2000; 56: 255-260
3. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. *CA: a cancer journal for clinicians* 2007; 57: 43-66
4. Malik G, Rojahn E, Ward MD, Gretzer MB, Partin AW, Semmes OJ, Veltri RW. SELDI protein profiling of dunning R-3327 derived cell lines: Identification of molecular markers of prostate cancer progression. *Prostate* 2007; 67: 1565-1575
5. Malik G, Ward MD, Gupta SK, Trosset MW, Grizzle WE, Adam BL, Diaz JI, Semmes OJ. Serum levels of an isoform of apolipoprotein A-II as a potential marker for prostate cancer. *Clin Cancer Res* 2005; 11: 1073-1085
6. Masters JRW. Clinical applications of expression profiling and proteomics in prostate cancer. *Anticancer Research* 2007; 27: 1273-1276
7. McDavid K, Lee J, Fulton JP, Tonita J, Thompson TD. Prostate cancer incidence and mortality rates and trends in the United States and Canada. Public Health Rep 2004; 119:174-86.

8. Nixon RG. Enhancing the specificity of prostate-specific antigen (PSA): An overview of PSA density, velocity and age-specific reference ranges. *British Journal of Urology* 1997; 79: 61-67
9. Petricoin EF, Ornstein DK, Liotta LA (2004) Clinical proteomics: Applications for prostate cancer biomarker discovery and detection. *Urologic oncology* **22:** 322-328
10. Pusch W, Kostrzewa M. Application of MALDI-TOF mass spectrometry in screening and diagnostic research. *Curr Pharm Des* 2005; 11: 2577-2591
11. Semmes OJ, Malik G, Ward M. Application of mass spectrometry to the discovery of biomarkers for detection of prostate cancer. *J. Cell. Biochem.* 2006 Jun 1; 98(3): 496-503.
12. Suckau D., Resemann A., Schuerenberg M., Hufnagel P., Franzen J., and Holle A. A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal. Bioanal. Chem.* 2003; 376: 952-965.
13. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, Feng Z, Parnes HL, Coltman CA, Jr. Assessing Prostate Cancer Risk: Results from the Prostate Cancer Prevention Trial. *J Natl Cancer Inst* 2006; 98: 529-534.
14. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, Minasian LM, Ford LG, Lippman SM, Crawford ED, Crowley JJ, Coltman CA, Jr. Prevalence of Prostate Cancer among Men with a Prostate-Specific Antigen Level <=4.0 ng per Milliliter. *N Engl J Med* 2004; 350: 2239-2246
15. Villanueva J, Philip J, Entenberg D, Chaparro CA, Tanwar MK, Holland EC, Tempst P. Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal Chem* 2004; 76: 1560-1570

16. Wright ME, Han DK, Aebersold R (2005) Mass spectrometry-based expression profiling of clinical prostate cancer. *Mol Cell Proteomics* **4:** 545-554
17. Xiao Z, Adam BL, Cazares LH, Clements MA, Davis JW, Schellhammer PF, Dalmasso EA, Wright GL, Jr. Quantitation of serum prostate-specific membrane antigen by a novel protein biochip immunoassay discriminates benign from malignant prostate disease. *Cancer Res* 2001; 61: 6029-6033
18. Xiao Z, Jiang X, Beckett ML, Wright GL. Generation of a baculovirus recombinant prostate specific mambrane antigen and its use in the development of a novel protein biochip quantitative immunoassy. *Protein Expression and Purification* 2000; 19: 12-21.

APPENDICES:

1. **Letter of Collaboration-** Dr. Ian M. Thompson Jr., Department of Urology, University of Texas Health Science Center, San Antonio, TX.
2. **Letter of Collaboration-** Dr. John E. Cornell, Department of Epidemiology and Biostatistics, University of Texas Health Science Center, San Antonio, TX.
3. **Published Abstract, IMPaCT 2007.**
4. **Manuscript Draft-** Title- "Identification of Candidate Prostate Cancer Biomarkers in Low PSA (< 4.0 ng/mL) Serum Samples Using MALDI-TOF and MALDI-TOF/TOF Mass Spectrometry".

The University of Texas
Health Science Center at San Antonio
Mail Code 7845
7703 Floyd Curl Drive
San Antonio, Texas 78229-3900

School of Medicine
Department of Urology
http://urology.uthscsa.edu/

(210) 567-5640
FAX: (210) 567-6868

Ian M. Thompson, Jr., M.D.
Professor and Chair
Henry B. and Edna Smith
Dielmann Memorial Chair
in Urologic Science
Glenda and Gary Woods
Distinguished Chair in
GU Oncology
Urologic Oncology

Stephen R. Kraus, M.D.
Associate Professor and Vice Chair
Hans H. Scharringhausen
Professor of Urology
Female Urology, Neurourology and Voiding Dysfunction

Joseph W. Basler, Ph.D., M.D.
Professor
Urology Residency Program Director
Thomas P. Ball, M.D. Residency
Educational Endowment
Urologic Oncology

Carlos E. Bermejo, M.D.
Assistant Professor
Urology Fellowship Program Director
Urologic Oncology

Jeffrey A. Leslie, M.D.
Assistant Professor
Pediatric Urology

Robert Marcovich, M.D.
Assistant Professor
Director of Endourology
Kidney Stones and Minimally Invasive Surgery

Dipen J. Parekh, M.D.
Associate Professor
Director of Robotic Surgery
Urologic Oncology

Dennis S. Peppas, M.D.
Professor
Pediatric Urology

Thomas A. Rozanski, M.D.
Professor
Reconstructive Urology

Sunil Sudarshan, M.D.
Assistant Professor
Urologic Oncology

Girard Seitter IV, FACHE
Administrator

December 11, 2007

Gunjan Malik, Ph.D.
Sr. Research Associate,
Division of Molecular Pathology
Cancer Therapy and Research Center's
The Institute for Drug Development
14960 Omicron Drive San Antonio,
TX 78245-3217

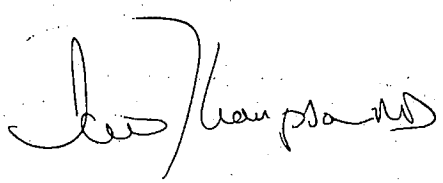
Dear Dr. Malik,

I would be happy to collaborate with you on the grant entitled- "Identification and Characterization of Prostate Cancer Associated Protein Biomarkers using High-throughput Mass Spectrometry". Your professional biomarker research expertise and the excellent facilities available at the Cancer Therapy and Research Center's Institute for Drug Development lend themselves extremely to the objective you have set forth for this project. This is a well-designed project that addresses the vital issues in cancer detection and diagnosis.

We have worked together on projects on SELDI EDRN Validation studies over the past years and have had discussions in laboratory meetings and conference calls. I have extensive experience in biomarker discovery and development and I totally support your proposal. We are in the process of developing a very sophisticated case-control sample set for prostate cancer that would be an interesting cohort to test on MALDI platform as well. I'll be more than happy to share this sample set with you per approved IRB protocols and MTA to assist in the second phase of your postdoctoral traineeship award. Your research approach has a great chance to generate better molecular markers of prostate cancer as well an overall understanding of the biology of cancer.

Good luck with your projects.

Sincerely,



Ian M. Thompson, M.D.
Professor and Chairman
Department of Urology

IMT/jjr

The University of Texas Health
Science Center at San Antonio
MC 7889, 7703 Floyd Curl Drive
San Antonio, Texas 78229-3900

Department of Epidemiology & Biostatistics

Phone: (210) 567-0836
Facsimile: (210) 567-0921

October 11, 2006

Gunjan Malik, Ph.D.
Sr. Research Associate,
Division of Molecular Pathology
Cancer Therapy and Research Center's
Institute for Drug Development
14960 Omicron Drive San Antonio,
TX 78245-3217

Dear Dr. Malik:

I would be happy to collaborate with you on the study entitled, "Identification and Characterization of Prostate Cancer Associated Protein Biomarkers using High-throughput Mass Spectrometry". It is a well-designed project that addresses the vital issues in cancer diagnostics.

I have extensive experience in biostatistical analysis of SELDI, MALDI, 2D-DIGE and Affymetrix data and I totally support your project. I will work with you on biostatistical analysis of proteomic data generated by your DOD-funded study and provide a proof-of-concept for further studies on well-designed clinical samples.

I'll be more than happy to also participate in the second phase of your project for validation of the cancer biomarkers in clinical samples using SELDI and/or MALDI-based immunoassays. Your research approach has a great prospect in generating new and novel molecular markers of prostate cancer as well as an overall understanding of the biology of cancer.

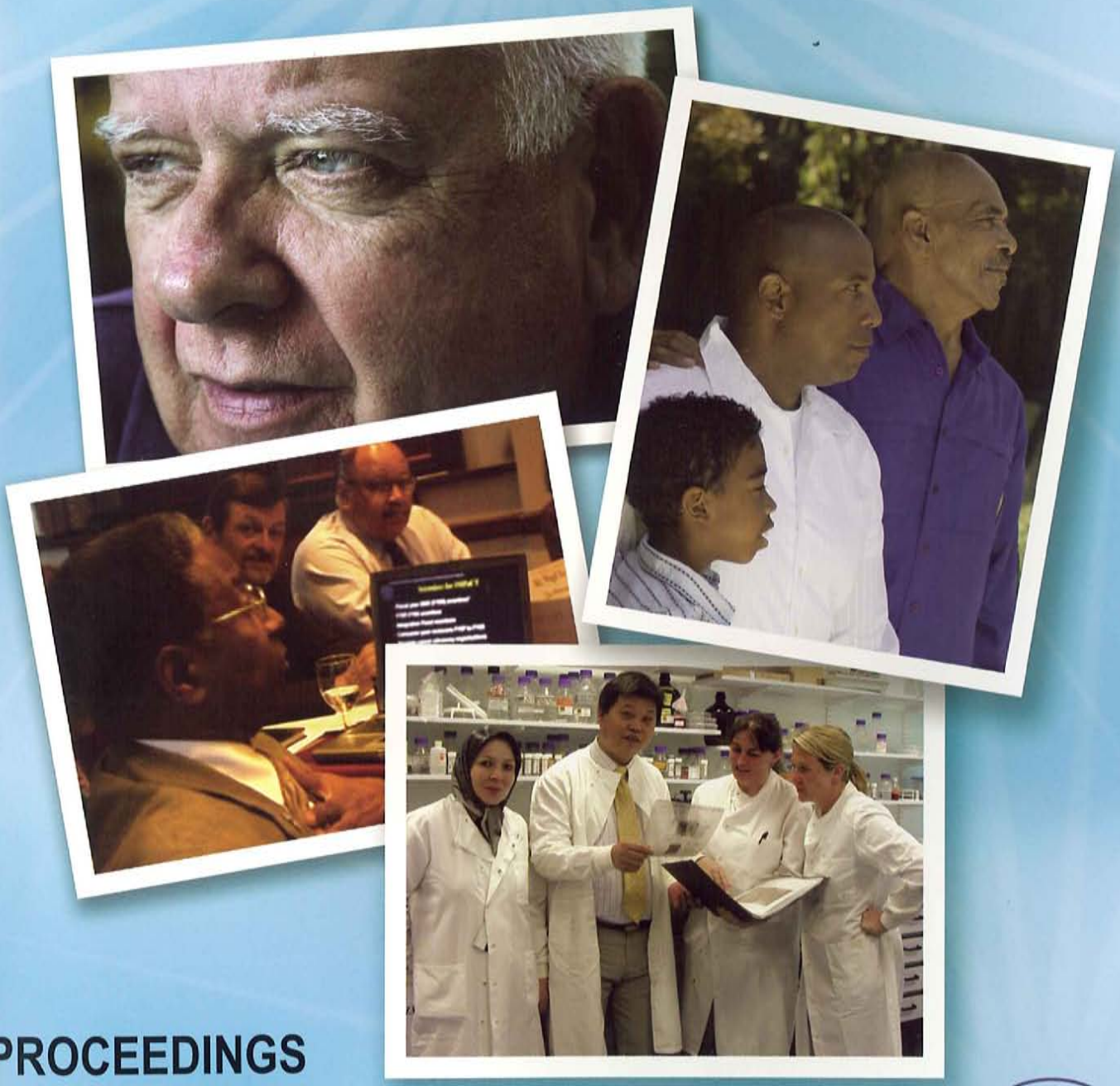
Good luck with your project.

Regards,

John E. Cornell, PhD
Professor
Department of Epidemiology and Biostatistics

Innovative Minds in Prostate Cancer Today

IMPaCT



PROCEEDINGS

September 5–8, 2007
Hyatt Regency Atlanta

IMPaCT

and poorly differentiated cancers. We found that pharmacological inhibition of MAOA resulted in alterations in prostate cancer cell growth.

Conclusions: The altered expression of genes associated with cancer cell differentiation provides functional insights into tumor phenotypes that influence tissue invasion, metastasis, and therapy resistance.

IMPaCT: The identification of genes and their cognate proteins that distinguish high-grade from low-grade carcinomas may be exploited to standardize cancer grading and possibly as circulating biomarkers capable of identifying aggressive disease. Further, several molecular features associated with high-grade prostate cancer involve aspects of cellular metabolism that can be inhibited using FDA-approved drugs (e.g., MAOA inhibitors).

The U.S. Army Medical Research and Materiel Command under W81XWH-05-1-0110 supported this work.

P26-10: No Association with Risk of Prostate Cancer for LDOC1 and SPANX-C Candidate Genes within the HPC-X Locus in a U.S. Study Population

Bradford Elmore, Joan Breyer, Kevin Bradley, Kate McReynolds, Jeffrey R. Smith, and Brian Yaspan
Vanderbilt University

We and others have previously undertaken linkage analysis in hereditary prostate cancer (HPC), typically in families with three or more cases, yielding significant loci difficult to confirm across study populations. Described HPC loci include Xq27-28 (HPC-X). Recent haplotype analysis of a Finnish founder population within the HPC-X locus has implicated *LDOC1* and *SPANX-C* as candidate prostate cancer genes, with a 150 kb region from markers D3S2390 to bG82i1.0 containing the critical region. *LDOC1* encodes a protein containing a leucine zipper-like motif and has been shown to be downregulated in some cancers. *SPANX-C* encodes a member of the sperm protein associated with the nucleus family, expressed solely in the testis and cancerous tissue. We sought to investigate *LDOC1* and *SPANX-C* as potential candidate genes in a study population of 597 U.S. Caucasian prostate cancer cases and 513 controls ascertained at Vanderbilt University. We screened for common polymorphisms and selected and genotyped 19 haplotype tagging SNPs (htSNPs) within and flanking *SPANX-C* and 1 htSNP within *LDOC1*. Analyses showed no evidence of increased prostate cancer risk due to genetic variation in *SPANX-C* or *LDOC1* in the study population. We are currently expanding our search for a hereditary prostate cancer gene in areas flanking this region.

IMPaCT: This project could potentially have an impact on prostate cancer detection by aiding identification of individuals at higher risk of developing prostate cancer. By identifying genetic determinants of prostate cancer risk, we can identify high risk individuals and suggest early prostate cancer screening thereby leading to increased early detection in high risk patients.

The U.S. Army Medical Research and Materiel Command under W81XWH-06-1-0057 supported this work.

P26-11: Expression of the Novel Survival Peptide, Humanin Protein Is Associated with Prostate Cancer Recurrence

Bingrong Liu, David Hwang, Hong Yu, Sheila Tze, Jonathan Said, David Seligson, Laura Cob, and Pinchas Cohen
University of California, Los Angeles

Humanin (HN) is a mitochondrial-encoded 24 amino acid polypeptide originally simultaneously discovered as a neuronal survival factor in Alzheimer's disease models and as an antagonistic binding partner for the pro-apoptotic molecules Bax and IGFBP-3. HN is secreted from cells and also acts by activating cell surface receptors and kinase cascades. We decided to investigate HNs potential significance in human prostate cancer by examining its in situ expression across a wide spectrum of primary tumors by tissue microarray analysis. Western immunoblotting was performed on frozen tissues from 10 cases of morphologically normal human prostate tissue and 10 samples of prostate cancer and revealed a single 3.5 kD band in both normal and cancer prostate samples with a fourfold stronger staining in prostate cancer than in normal prostates ($p<0.001$). Similarly, TRAMP tumors stained 4–5 times stronger than normal 20-week old mouse prostates. Immunohistochemistry was performed on tissue microarrays constructed from paraffin embedded primary prostate cancer specimens from 226 hormone naïve patients who underwent radical retropubic prostatectomy. In total, 979 tissue microarray spots were informative, including morphologically normal prostate (NL; n=257), benign prostatic hyperplasia (BPH; n=107), prostatic intraepithelial neoplasia (PIN; n=41) and invasive prostate cancer (cancer; n=574). Both cytoplasmic and nuclear HN expression was scored in a semi-quantitative fashion using an integrated intensity measure (0.0–3.0) and positivity (0%–100%), respectively. The protein expression distribution was examined across the spectrum of epithelial tissues and its association with standard clinicopathological covariates and tumor recurrence was examined in 184 outcome and marker-informative patients. HN expression was low overall in prostate tissues. Ninty-seven percent of spots had negative to weak cytoplasmic staining (<1.0), and 92% of spots had infrequent nuclear staining (<25%). Nonetheless, the mean cytoplasmic HN expression was significantly higher in cancer (intensity = 0.11) compared to normal (intensity = 0.045; p=0028), and the mean nuclear HN expression was also significantly higher in cancer (10.48% positive) compared to normal tissue (2.24% positive; $p<0.0001$). Nuclear HN expression was a significant and negative prognosticator in low-grade (Gleasons Score 2–6) prostate cancers when used as either a continuous variable (Cox Proportional Hazards p=0.006), or dichotomized variable cut at 15% positive (p=0.01). However, HN was not an independent predictor of tumor recurrence in multivariate analysis in this patient substrata, including seminal vesicle and capsular involvement, and preoperative PSA as covariates. We conclude that HN is expressed at higher levels in prostate cancers compared to matched normal tissues. High nuclear HN expression is strongly associated with an increased risk of tumor recurrence. This is the first report examining HN in prostate cancer and our findings suggest that it may be a prognostic marker and therapeutic target and is most informative in patients with low-grade tumors, providing further stratification of this category as a potential guide for clinical follow-up and clinical trials.

The U.S. Army Medical Research and Materiel Command under W81XWH-06-1-0211 supported this work.

P26-12: Identification and Characterization of Prostate Cancer Associated Protein Biomarkers Using High-throughput Mass Spectrometry

Lisa H. Cazares[1], Shamina G. Mitchell[1], Mary Ann Clements[1], Tarek Kandil[1], Brian Main[1], O. John Semmes[1], Jose I. Diaz[2], and Gunjan Malik[2]
[1]Eastern Virginia Medical School
[2]University of Texas Health Science Center at San Antonio

Due to the high prevalence of prostate cancer (PCa), there has been a large-scale search for potential biomarkers useful in the early detection and prognosis of PCa. Although there have been a lot of biomarker discoveries over past few years, validation of their clinical utility has not been accomplished. One of the reasons for the lack of their successful clinical utilization is that no single marker can accurately reflect the complex phenotypic changes associated with development of cancer. The development of high-throughput methods which are able to analyze large segments of the proteome promise to facilitate the identification of multiple protein panels for cancer diagnostics. Our laboratory pioncered the application of Surface-Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF MS) for separation and analysis of complex mixtures of proteins. We have also developed improved combinatorial approaches for rapid identification of the protein biomarkers using LC-MS/MS and MALDI-TOF/TOF and have further refined mass spectrometry-assisted immunoassays for detection and quantification of potential biomarkers in body fluids.

The objective of this study is to use state-of-the-art MALDI-TOF/TOF and LC-MS/MS to identify and characterize prostate cancer-associated protein biomarkers to aid in the development of improved clinical assays for early detection/diagnosis and prognosis of prostate cancer. With the use of high-throughput affinity chromatography with front end robotics, to enhance throughput and reproducibility and Tandem Mass Spectrometers, the project will result in (1) discovery of the proteins differentially expressed between cancer and non-cancer using MALDI-TOF MS, (2) identification of the proteins with diagnostic potential, and (3) development of an MS-assisted immunoassay for PCa diagnostics.

We utilized sera drawn from cancer patients with marginal clinical symptoms (PSA <4.0 ng/mL), processed and analyzed using MALDI-TOF system. Protein profiles, generated by MALDI-TOF-MS, were analyzed using Ultraflex™ software. Comparative analysis of the protein profiles identified

the size (mass/charge) of the proteins differentially expressed between sera from normal and diseased cases. Prostate cancer associated protein biomarkers, identified using MALDI, can now be partially purified by HPLC and FPLC techniques and identified using tandem mass spectrometry for development of new/novel clinical biomarkers for PCa detection/diagnosis.

IMPaCT: This project could have a large impact on Prostate Cancer through more accurate identification of PCa either at clinical presentation or post biopsy and thus reducing the number of needless biopsies and assessing the risk of successive biopsies in high-risk patients.

The U.S. Army Medical Research and Materiel Command under W81XWH-06-1-0111 supported this work.

P26-13: Telomere Attrition of Isolated High-grade Prostatic Intraepithelial Neoplasia and Surrounding Stroma Is Predictive of Prostate Cancer

Anthony M. Joshua[1], Bisera Vukovic[1], Ilan Braude[1], Sundus Hussien[2], Maria Zielenska[3], John Srigley[2], Andrew Evans[1], and Jeremy A. Squire[1]
[1]University Health Network, Toronto
[2]Credit Valley Hospital
[3]Hospital for Sick Children

Background and Objectives: The causes of the early genomic events underlying the development of prostate cancer (CaP) remain unclear. Onset of chromosomal instability is likely to facilitate the formation of crucial genomic aberrations in both the precursor lesion, high-grade prostatic intraepithelial neoplasia (HPIN) and in CaP. Instability generated by telomere attrition is one potential mechanism that could initiate chromosomal rearrangements.

Methodologies: In this study the normalised telomere length variation was examined in a cohort of 68 men without CaP, who had HPIN only on prostatic biopsies. Quantitative fluorescence in-situ hybridisation (Q-FISH) was performed using a telomeric peptide nucleic acid (PNA) probe and pan-centromeric control PNA probes on one biopsy from each man.

Results: Multiple significant associations between telomere attrition and an eventual diagnosis of CaP in the HPIN as well as in the surrounding stroma were found. Logistic regression analysis revealed significant associations ($p<0.05$) between telomere length and outcome to CaP. Kaplan-Meier analysis of telomere length demonstrated significant increased risk for the development of cancer with short telomeres in surrounding stroma ($p=0.035$, HR=2.12, 95% CI 0.231–0.956), and a trend for HPIN itself ($p=0.126$, HR=1.72, 95% CI 0.287–1.168). Cox regression analysis also demonstrated significance between the time from the original biopsy to the diagnosis of cancer and telomere length in HPIN and in the surrounding stroma. Additionally, there appeared to be a trend suggesting that sites of the prostate in which cancer eventually developed had much shorter telomere lengths than those that were free of cancer on follow-up.

Conclusions: These analyses suggest that telomere length is a risk factor for the eventual diagnosis of prostate cancer in men who have HPIN only on their prostate biopsies. These results lend support to the hypothesis that telomere attrition in prostatic preneoplasia may be fundamental to the generation of the chromosomal instability and to the emergence of CaP. The finding of concordant telomere length shortening in surrounding stroma adjacent to foci of HPIN maybe somatic or constitutional in nature and raise important hypotheses about the nature of prostatic carcinogenesis with the potential for senescence related stromal factors to contribute to CaP progression. Further work is examining both telomere length in various prostate pathologies and the consequence of telomere dysfunction on the DNA damage response.

IMPaCT: These results support a critical role for telomere dysfunction in prostatic carcinogenesis, with both clinical and scientific consequences. Patients with HPIN may be able to be risk stratified on the basis of telomere length for future occurrence of CaP whilst insights from this research applies broadly across phenomena relating to chromosomal instability, tumour-stromal interaction and the field carcinogenesis effect in prostatic carcinogenesis.

The U.S. Army Medical Research and Materiel Command under W81XWH-05-1-0619 supported this work.

P26-14: Serum Glycan Profiling as a Prognostic Indicator for Prostate Cancer

Crystal Kirmiz, Ruth Vinall, David Rocke, Carlito Lebrilla, Ralph deVere White, and Suzanne Miyamoto
University of California, Davis

Prostate cancer is the second leading cause of cancer death among men. Although a screening test that measures prostate specific antigen (PSA) is currently available for detecting prostate cancer, elevated PSA values are also possible for benign prostatic hyperplasia, a non-cancerous condition, which then makes it necessary to obtain a surgical biopsy. Even after the biopsy is performed and the patient is diagnosed with prostate cancer, determining the type of treatment (surgery, watchful waiting, or radiation therapy) is often difficult. During the development of prostate cancer, tumor cells change their proteins and glycosylation of proteins. Glycosylation is the attachment of sugar groups (glycans) to extracellular proteins, which then influence growth, motility, and immune surveillance of the tumor. These glycosylation changes can be correlated with increasing tumor burden and poor prognosis. We have developed methods to detect the presence of aberrant glycans in serum from cancer patients using sensitive analytical methods. Only a small volume of serum is needed for this method. The glycans are chemically cleaved from their protein core and separated into neutral and anionic glycans by solid phase extraction. The resultant fractions are then analyzed by high resolution Fourier transform ion cyclotron resonance (FTICR) mass spectrometry (MS) and a glycan profile is produced from each patient serum. The identity of a mass as a glycan and further structural information is obtained by tandem mass spectrometry using infrared multiphoton dissociation (IRMPD). Using funds provided by the DOD Prostate Cancer Research Program FY2005 Concept Award, we were able to demonstrate our method on conditioned media from three prostate cancer tumor cell lines, LNCap, PC3 and SaOS, each of which contained either a vector control or had been stably transfected with a gain-of-function p53 mutant (R273H). Glycan profiles were prepared from multiple samples from each cell line. Each cell line produced a distinct glycan profile that changed when transfected with the p53 R273H gain of function mutant. The LNCaP and PC3 tumor cell lines produced similar glycan profiles, whereas the Saos-2 cells produced a different profile. The method was then optimized for analysis of patient serum samples. Two groups of prostate cancer patient sera are being tested. One group of serum samples came from prostate cancer patients after radical prostatectomy and the other group came from cancer patients undergoing "watchful waiting." Glycan profiles have been prepared from these samples and manual analysis shows measurable differences between the two groups of samples. Further bioinformatic and statistical analysis is being used to find additional differences and similarities between the glycan masses of each profile. In addition to the β elimination method another method of cleaving glycans PGNase enzymatic cleavage is being used to release the N-linked glycans. The results of this analysis will also be presented. With our results we plan to submit proposals for future funding to the CDMRP and NIH. Results from this translational study have the potential for relevant biomarkers of prostate cancer that can be used clinically to assist in the diagnosis and prognosis of prostate cancer and help to guide treatment decisions.

The U.S. Army Medical Research and Materiel Command under W81XWH-06-1-0011 supported this work.

P26-15: Evaluation of Genomic Instability by Methylation Status in the Abnormal Prostate

Kimberly Butler, Jeffrey Griffith, Christina M Haaland-Pullus, Christopher Heaphy, and Marco Bisoffi
University of New Mexico, Albuquerque

Prostate cancer is the second most common cause of cancer related death in men after lung cancer, with incidence of prostate cancer increasing significantly with advanced age. It is currently accepted that tumorigenesis is that a multi-step process where there is accumulation of genetic and epigenetic changes that alter the normal regulatory mechanisms controlling cellula proliferation. However, not enough is yet known about the processes o tumorigenesis and disease progression, creating limitations in detection treatment, and prevention of this cancer. This study is designed to look fo better methods of detection and prognostic markers regarding prostate cance to reduce the risk of mortality associated with current treatment modalities.

Identification of Candidate Prostate Cancer Biomarkers in Low PSA (< 4.0 ng/mL) Serum Samples Using MALDI-TOF and MALDI-TOF/TOF Mass Spectrometry

[1]**Gunjan Malik**, [2]Lisa H. Cazares, [1]Saurabh K. Gupta, [3]John E. Cornell, [4]Kali. H. Makedou, [2]Shamina G. Mitchell, [2]MaryAnn Clements, [2]Tarek O. Kandil, [2]Brian P. Main, [2]Richard R. Drake, [2]O. John Semmes and [1]Jose I. Diaz

[1]Department of Molecular Pathology, Cancer Therapy and Research Center's Institute for Drug Development at the San Antonio Cancer Institute, San Antonio TX

[2]Center for Biomedical Proteomics, Eastern Virginia Medical School, Norfolk VA

[3]Department of Epidemiology and Biostatistics, University of Texas Health Science Center, San Antonio TX

[4]Department of Biochemistry, School of Medicine, Aristotle University of Thessaloniki, Greece

Running Title: MALDI-TOF/TOF mass spectrometry of low PSA prostate cancer.

Keywords: MALDI, Biomarker, Identification, Prostate Cancer.

Grant Support: This study was supported by the Department of Defense (DOD) Prostate Cancer Research Program (PCRP) of the Office of the Congressionally Directed Medical Research Programs (CDMRP) (# W81XWH-06-1-0111; P.I. Gunjan Malik, Ph.D.)

Address of Correspondence:

CTRC\Institute for Drug Development

14960 Omicron Drive, San Antonio, TX- 78245-3217

Ph: (210) 450-3800 FAX- (210) 450-3939 Email- gmalik@idd.org

ABSTRACT:

Purpose: Prostate Specific Antigen (PSA) testing has tremendously increased the detection of early-stage prostate cancer (PCa). However, a PSA value higher than 4ng/mL warrants a biopsy that often indicates benign disease. On the other hand, recent assessments reveal an equally elevated risk (20-25% incidence) of PCa among men with serum PSA levels from 2.5 - 4.0 ng/mL. Our objective was to determine if serum protein-expression profiles could be used to differentiate between benign and malignant prostate cancer in biopsy proven case (biopsy positive) and control (biopsy negative) patients with low serum PSA (< 4.0 ng/mL).

Experimental Design: Serum was collected ± four weeks from the date of biopsy from prostate patients with a positive (CA) or at least three negative biopsies (NO). Sera were incubated in duplicates with- immobilized metal ion affinity magnetic beads charged with copper (IMAC-Cu); and weak cation exchange magnetic beads (WCX) using the ClinProt automated workstation. Samples were analyzed on matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) instrument (Bruker Daltonics) in both linear and reflector modes. Spectra were processed and analyzed using ClinProTools 2.0 software (Bruker Daltonics), and classifications determined using genetic-clustering and AdaBoost algorithms. MALDI-LIFT-TOF/TOF was applied to identify protein/peptide peaks of strongest significance using Biotools 2.2 (Bruker Daltonics).

Results: Sera from a total of 174 subjects were selected to form cancer (CA; n = 68) and non-cancer (NO; n = 106) cohorts. The MALDI-TOF mass spectrometry yielded a total of 448 peaks, with 65 peaks expressed differentially ($p < 0.05$) between the cancer and non-cancer cohorts. The AdaBoost algorithm generated a sensitivity of 73.5% and

specificity of 93.8% with area under the ROC curve of 0.94. MALDI-LIFT-TOF/TOF spectrometry identified some of the peaks of statistical significance which were verified by Western blotting.

Conclusions: MALDI-TOF protein-expression profiles generated from sera (PSA < 4 ng/mL) could be used to distinguish between cancer and non-cancer cases of prostate disease.

BACKGROUND: The number of individuals affected by cancer continues to rise as our life expectancy increases. A total of 1,444,920 new cancer cases and 559,650 deaths for cancers are projected to occur in the United States in 2007. In fact, the incidence for urological cancers such as that of the prostate continues to climb with each successive year of life, making it the most common cancer in men. Prostate cancer (PCa) is the most frequently diagnosed cancer in men. With an estimated 27,050 deaths from PCa, it is a leading cause of cancer death in men. More than 90% of all PCa are discovered in the local and regional stages with their 5-year survival rate reaching almost 100%. However, the survival rate drops to 33% when PCa has spread to distant sites. Approximately 40,000 men die each year with PCa metastasis (Jemal et al, 2007). Despite the long-time use of a 4.0 ng/mL cutoff for a 'normal' Prostate Specific Antigen (PSA) levels in the blood, it has been acknowledged that only about 25% of men with such an elevated value will be found to have cancer at prostate biopsy. Because of this, three quarters of men with an elevated PSA who have a biopsy undergo the procedure unnecessarily. Recent data from the PCPT trial suggest that the risk of PCa is equally elevated (20-25% incidence) even among men with serum PSA levels from 2.5 to 4.0 ng/mL. Additional indications for prostate biopsy include a rising PSA, an abnormal DRE, or lower PSA with other risk factors such as history of PCa and/or prior (negative) biopsies.

The detection of such operable cancers earlier, the identification of indolent cancer and the avoidance of unnecessary biopsies are all promises of better molecular-based early detection efforts. High-throughput expression profiling approaches hold a tremendous potential for identifying biomarkers which could be helpful in detection, diagnosis and targeted therapy of cancer. The tremendous advances that have been made in high-

throughput "omics" technologies (e.g., genomics, transcriptomics, proteomics and metabolomics) are providing the most comprehensive means to identify candidate molecular markers of cancer (Semmes et al, 2006; Wulfkuhle et al, 2004; Zhang et al, 2007). The potential impact of these multifaceted discovery technologies on cancer diagnostics and prognostics can be realized via two complementary but separate directions. The first is the utility of the unique “fingerprint” pattern derived from the protein expression data. The second is the discovery of actual protein/peptide biomarkers that can be subsequently utilized in an immunoassay or other multiplexed display array platforms. Proteomic techniques aimed at biomarker discovery have been centered on identification of differentially expressed proteins following gel or liquid chromatographic separation. The candidate biomarker is then evaluated by immunoassay for population-wide sensitivity and specificity at detection. This two step approach is proven to be effective and has been greatly enhanced by the sequencing of the human genome and concomitant improvements in mass spectroscopy. 2D-gel analysis has been the proteomic tool of choice, with systems now routinely analyzing 10s of gels simultaneously. However, in addition to the need for high-throughput, there is a tremendous need for improved ability to “mine” the full depth of the proteome. Methodology that can accommodate higher-throughput with the ability to observe high volume of protein events are needed to advance clinical proteomics. Currently, many systems that couple robotic handling of samples in the front-end to a MALDI-TOF mass spectrometer are being evaluated for clinical utility.

The current study aimed at establishing the clinical utility of high-throughput MALDI-TOF approach to protein profiling for specific early detection objectives in prostate

cancer in men with PSA <4.0 ng/mL where there is still a significant percentage of PCa left undetected (~25%). We applied the MALDI-LIFT TOF/TOF approach to the characterization and sequence identification of potential biomarkers for detecting prostate cancer. We examined serum samples as readily available, relatively non-invasive source of cancer biomarkers. Identification of these cancer biomarkers will assist in development of better non-invasive diagnostic tools for prostate cancer. Further understanding of these biomarkers and their functional aspects may also eventually help in better perception of the biology of cancer and lead to better therapeutic targets. Development of new and novel biomarkers in this clinical *gray area* may also prove useful for contributing to the PSA test by complementing this marker in the range where PSA has *failed* to detect cancer.

METHODS:

***STUDY DESIGN*:** Serum samples were collected from biopsy proven cases and controls within ± 16 weeks from the date of biopsy. A total of 68 cases and 106 controls were collected for this pilot study to allow an initial review of the sensitivity and specificity of the test. All the "normal" controls (biopsy negative; NO) and cases (biopsy positive; CA) had a PSA level from 0 - 4 ng/mL. The samples were collected and processed using the standard protocols for serum collection and stored at -80°C. All samples were obtained from properly consented patients through the institutional review board approved protocols.

***MALDI-TOF MS*:** The cancer and control serum samples were randomized over 96-well plates or bioprocessors (BP) along with randomly placed quality control serum (QC) as reference controls. Serum samples were assayed randomly with two different types of magnetic beads (Bruker Daltonics) with different binding affinities- MB-IMAC-Cu: Immobilized metal affinity charged with copper and MB-WCX: Weak cation exchange magnetic beads. Binding of serum samples to the magnetic beads was performed in duplicates, according to the manufacturer's recommendations on a ClinProt Automated Laboratory Workstation. Briefly, for each analysis, 20 μL serum was incubated with 10 uL magnetic beads as per manufacturer's instructions (Bruker Daltonics). Unbound proteins were discarded, and each sample washed three times in binding buffer. Bound proteins were eluted as per manufacturer's instructions, and spotted in duplicate on an AnchorChip sample target platform (384 spots), mixed 1:10 with α-cyano-4-hydroxycinnamic acid (CHCA in an acetone:ethanol mixture of 1:2). To run the samples in reflector mode, eluted proteins were mixed 1:5 in CHCA matrix with 0.25% TFA.

Samples were run in both linear (0-100,000 *m/z*) and reflector (0-10,000 *m/z*) mode on an Ultraflex III matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) instrument (Bruker Daltonics) controlled by the Flex Control 3.0 software package. Peptide and Protein Standards (Bruker Daltonik GmbH, Germany) were used for calibration of the respective mass range. Flex Analysis 3.0 software was used to assess the spectra and ClinProt 2.0 software was used for normalization of spectra (using total ion current), baseline subtraction, calibration, peak labeling (mass-to-charge ratio or *m/z* values) and peak alignment (in the ± 0.2% *m/z* window).

K-nearest neighbor genetic algorithm and support vector machine algorithm, contained in the software suite were used to select the protein/peptide peaks with most statistically significant differences in the two groups analyzed. After each model was generated, internal cross-validation was applied within the software to determine the sensitivity and specificity of the classifications.

***PROTEIN IDENTIFICATION*:** To identify some of the protein peaks ($p < 0.05$) generated by the MALDI-TOF, MALDI-LIFT-TOF/TOF spectrometry (Suckau D et al. *Anal Bioanal. Chem.* 2003 Aug; 376(7):952-65) was applied to the eluates from the magnetic beads in the presence of CHCA matrix. Fragment ion spectra were first analyzed with FlexAnalysis 3.0 (Bruker Daltonik GmbH, Germany). Peptide mass fingerprints generated by the MALDI-LIFT approach were used for MASCOT (Matrix Science, London, UK) search employing Biotools 2.2 (Bruker Daltonics).

***BIOMARKER VALIDATION*:** To validate the identity of the proteins identified using the MALDI-LIFT approach, Western analysis was performed on 6uL of the eluted proteins derived from the automated processing of serum samples on ClinProt robot using

IMAC and WCX magnetic beads. Case and control eluates were run parallel on 4-12% pre-cast Criterion gels (Bio-Rad Laboratories, Inc.) and transferred to PVDF membrane at 400mA for 50min. using the trans-blot semi-dry transfer cells. Primary antibodies were obtained from Abcam, Inc. and HRP-conjugated secondary antibodies were obtained from Bio-Rad Laboratories, Inc.

RESULTS:

Sera from prostate cancer (n= 68) and controls (n= 106) were processed on IMAC-Cu and WCX magnetic beads using the robotic ClinProt workstation. Bruker Ultraflex MALDI-TOF MS spectra were acquired from each set of eluates in both linear and reflector mode (**Figure 1**). ClinProt 2.0 analysis yielded a total of 448 peaks, with 65 peaks expressed differentially ($p < 0.05$) between the cancer and control cohorts (**Table 1**). Most of the protein peaks with a p-value < 0.05 ranged from 2000-30,000 *m/z* in the linear mode and 1000-3000 *m/z* in the reflector mode (**Table 2**). Classification models were generated using Genetic algorithm and Support Vector Machine via ClinProt 2.0 software for each set, yielding the best sensitivity of 71.3% and specificity of 82.7% (**Table 3**). We then applied the AdaBoost algorithm (Qu et al. *Clinical Chemistry* 2002; 48: 1835-1843) with the J48 algorithm for growing and pruning decision trees (Weka 3.5.6) to this set and used a 10-fold internal cross-validation method. The strongest classification was obtained using this combination generating **73.5% sensitivity** and **93.8% specificity** for classifying cancer and non-cancer cases.

All peaks above an intensity threshold of 500 relative intensity units and separated from neighboring peaks by at least 10 Da were subjected to MS/MS analysis. In total 55 peaks from the ConA fractions, 43 peaks from the WGA fractions and 37 peaks from the

boronic acid fractions were analysed by MS/MS. Examples for the MS/MS spectra of a tryptic peptide of Histidine-rich glycoprotein (1124.582 Da) and of a tryptic peptide of Serum Amyloid P component (1811.971 Da) are given in Fig. 5. The resulting fragment spectra were submitted to MASCOT for database search with the objective to identify the corresppponding proteins. For the ConA fraction 45 MS/MS spectra, for the WGA fraction 21 spectra and for the boronic acid fraction 17 spectra led to significant hits revealing 12, 10 and 10 different maternal proteins, respectively. The binding profiles of the different beads comprised different and also identical proteins. Alpha-2-macroglobulin, Ceruloplasmin and Histidine-rich glycoprotein were bound by ConA and WGA. Kininogen was found in the WGA and the boronic acid fraction and Complement C1q was detected by ConA and boronic acid.

For protein/peptide identification, MALDI-LIFT-TOF/TOF spectrometry was performed on the eluates from the magnetic beads in the presence of CHCA matrix. Out of a total of 22 peaks or “parent ions” from MB-WCX, 19 peaks could be identified. However, out of the 10 peaks from MB-IMAC, none of the proteins could be identified probably due to some kind of post-translational modification(s) and/or processing (**Figure 3**). One of the proteins identified from more than one parent ion was **Complement C3**. Western Analysis of some of the case and control eluates from MB-WCX using anti-complement C3 antibodies yielded a 72kDa protein band differentially expressed between the case and control groups (**Figure 3C**). A notable finding was the observation of an 8.9K *m/z* protein peak in the MB-IMAC-Cu data set run in the linear mode (**Figure 4**). Our previous studies using SELDI-TOF MS have reported the up-regulation of an 8.9kDa isoform of Apolipoprotein A-II in PCa even in the low PSA samples on IMAC-Cu ProteinChips

(Malik G et al. *Clin. Cancer Res.* 2005 Feb 1;11(3):1073-85). Studies are underway to identify this 8.9K *m/z* protein peak displaying a consistent overexpression in PCa in a similar data set on MALDI.

CONCLUSION

MALDI-TOF protein-expression profiles generated from prostate cancer sera could be used to distinguish cancer from non-cancer sets with a relatively good sensitivity and specificity. “On-the-flight” protein identification using the MALDI-LIFT technology can provide insight into the “fingerprint” profiles and provide stronger tests for cancer detection.

Author Contributions

Study conception and design: Malik, Semmes, Diaz

Acquisition of data: Malik, Cazares

Analysis and interpretation of data: Malik, Cazares, Cornell

Drafting of manuscript: Malik, Semmes, Diaz

Table 1. Number of protein peaks detected by MALDI-TOF MS

Sample Processing	Total number of peaks*	Number of peaks with a PTT§ <0.05
IMAC-Linear	131	16
IMAC-Reflector	43	10
WCX-Linear	61	21
WCX-Reflector	213	18
TOTAL	**448**	**65**

*Total number of "peaks" generated by Flex Analysis 3.0

§Number of "peaks" with p-value of Student's t-test <0.05

Table 2. Protein peaks with a p-value <0.05 generated by Flex Analysis 3.0 (Bruker Daltonics)

IMAC Linear		IMAC Reflector		WCX Linear		WCX Reflector	
Mass	**PTT***	**Mass**	**PTT***	**Mass**	**PTT***	**Mass**	**PTT***
2192.48	0.0103	1349.05	0.005	2195.37	0.0085	929.3	0.0294
2607.69	0.0051	1451.03	0.005	2656.6	0.0006	1046.33	0.0241
2644.87	0.0002	1779.35	0.0073	2742.76	0.0074	1061.22	0.0207
2714.03	0.0532	1866.43	0.0073	2853.43	0.0042	1450.6	0.0294
2734.83	0.0466	2022.59	0.005	3050.57	0.0321	1692.67	0.0223
2918.37	0.0083	2082.41	0.0545	3809.49	0.0321	1779.75	0.0207
2937.58	0.011	2210.53	0.005	4461.82	0.023	1866.83	0.0207
3227.49	0.0002	2645.6	0.0164	4790.22	0.0002	1888.84	0.0294
3248.8	0.0151	2660.88	0.0181	5470.27	0.0532	2007.16	0.0024
4199.45	0.0076	2933.29	0.0288	5904.52	0.0042	2022.99	0.0024
4395.02	0.0002	3225.8	0.043	6435.99	0.0265	2211.06	0.0167
5037.71	0.0076			6636.29	0.0407	2239.07	0.0031
5044.66	0.038			7460.2	0.035	2367.25	0.0155
5056.68	0.0076			7771.69	0.0265	2624.42	0.0294
5899.29	0.0003			8140.44	0.0301	2646.18	0.0033
7764.09	0.0151			8925.57	0.0042	2661.44	0.0167
8921.76	0.0046			10272.77	0.023	2791.17	0.0277
				10665.39	0.0093	3242.73	0.0167
				12602.4	0.0172		
				14047.26	0.0321		
				15178.77	0.0266		
				28010.96	0.0353		

*p-value of Student's t-test

Table 3. Classification of the case and control sets using various algorithms

	Model Generation			*Internal cross-validation*		
	CANCER	**CONTROL**	**OVERALL**	**CANCER**	**CONTROL**	**OVERALL**
IMAC-Lin. (GA)	100 %	100 %	100 %	70.36 %	77.22 %	**73.79 %**
IMAC-Ref. (GA)	100 %	100 %	100 %	69.5 %	80.95 %	**75.23 %**
WCX-Lin. (GA)	100 %	100 %	100 %	63.42 %	74.83 %	**69.13 %**
WCX-Ref. (GA)	100 %	99.07 %	99.53 %	56.98 %	71.72 %	**64.35 %**
WCX-Ref. (SVM)	100 %	99.02 %	99.51 %	71.26 %	82.74 %	**77.00 %**
WCX-Lin. (AdaBoost)	**-**	**-**	**-**	**73.5 % (Sensitivity)**	**93.8 % (Specificity)**	**85.9 % (AUC 0.94)**

GA- k-nearest neighbor genetic algorithm (ClinProt 2.0, Bruker Daltonics).; SVM- Support Vector Machine algorithm (ClinProt 2.0, Bruker Daltonics); AdaBoost- Boosting algorithm (Qu et al. 2002) with a J48 algorithm for growing and pruning decision trees (Weka 3.5.6).

FIGURE LEGENDS

Figure 1. Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Analysis: Left: The calculated average spectra for the case (red) and control (green) classes is shown for the 1000 to 35,000 *m/z* range for the linear mode and 1000 to 3500 *m/z* range for the reflector mode. The x-axis records the *m/z* value (mass-to-charge ratio), the y-axis is the peak intensity in arbitrary units (arb. u.). The plot is drawn on a unique scale independent of the peak intensity scale. **Right:** Heat map overview of MALDI spectra. Normalized peak intensities for each of the 348 spectra generated in each data set is shown for the 1500 to 10,000 *m/z* range for the linear mode and 500 to 3500 *m/z* range for the reflector mode. The green arrowheads indicate peaks overexpressed in the control samples and red arrowheads indicate peaks overexpressed in the cancer samples.

Figure 2. Representative matrix-assisted laser desorption/ionization time-of-flight spectra. Two representative spectra, one from each cohort (cancer CA and control CO) were selected randomly from the ~1500 total spectra generated in the analysis of the clinical samples after IMAC and WCX magnetic bead enrichment. Mass range of 2,000 to 20,000 *m/z* (linear mode) and 1000- 3000 *m/z* (reflector mode) is shown. The arrows indicate peaks displaying differential expression in the case vs. the control set.

Figure 3. MALDI-LIFT TOF/TOF MS. A. Peak distribution plot displaying the areas of the respective peaks in each single spectrum of the two classes as separate values. Peak distribution of each sample analyzed in the 1000 to 2200 *m/z* range in MB-WCX in the reflector mode are shown [green circles- controls; red crosses- cancer]. The peaks at 1061, 1779, 1866 and 2022 *m/z* are shown with their respective p-values in parenthesis.

B. All the protein peaks displayed in panel A were identified by **MALDI-LIFT TOF/TOF** as fragments of **Complement Component C3**. Fragment ions observed in an MS/MS spectrum of 1779.75 *m/z* protein peak is shown in B, marking the fragment ions with the Biemann nomenclature. **C.** Western Analysis of some of the MB-WCX eluates from cases (lanes 3 and 4) and controls (lanes 1 and 2) using anti-Complement C3 antibodies (Abcam, Inc.) detected a 72kDa protein overexpressed in cancer samples validating its identity.

Examples of two MS/MS spectra acquired on an autoflex II TOF/TOF. (A) MS/MS spectrum of the peptide peak *m*/*z* = 1124.5 Da representing the peptide aa 44–52 of Histidine rich glycoprotein. (B) MS/MS spectrum of the peptide peak *m*/*z* = 1811.9 Da representing the peptide aa 150–165 of Serum Amyloid P component

Figure 4. Panel A: Average spectrum view with a close-up look at the 7500-9500 *m/z* range. Arrow points to the 8.9K *m/z* protein peak overexpressed in cancer with its p-value in parenthesis. Respective heat map of the same *m/z* range is shown in **Panel B** with the 8.9K *m/z* peak highlighted in the box. **Panel C** [*reprinted from* Malik G et al. (*Clin. Cancer Res.* 2005 Feb 1;11(3):1073-85.)] displays the expression of 8.9K *m/z* peak in prostate serum samples with low PSA in 40 cases and 154 control samples as observed in a previous study using SELDI-TOF MS on IMAC-Cu2+.

FIGURES

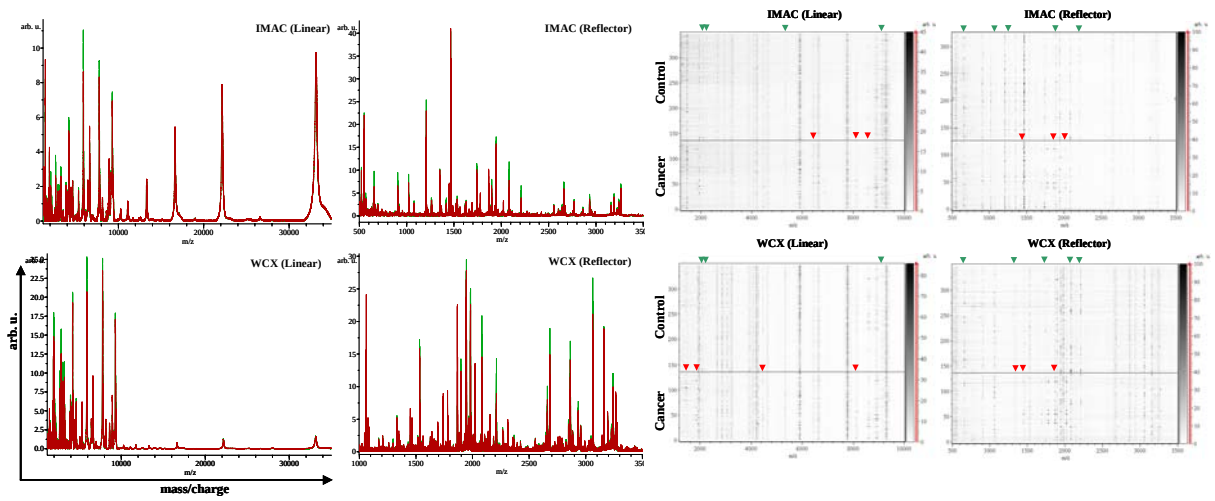


Figure 1.

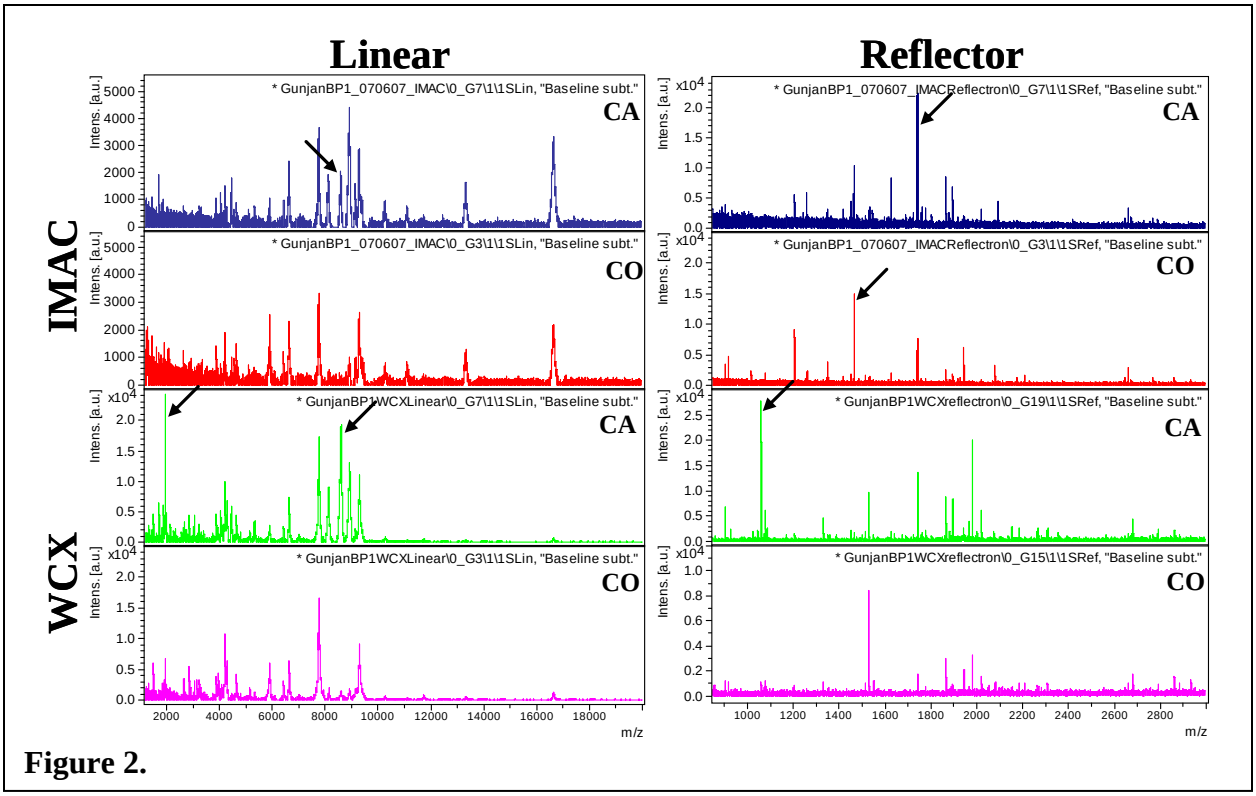


Figure 2.

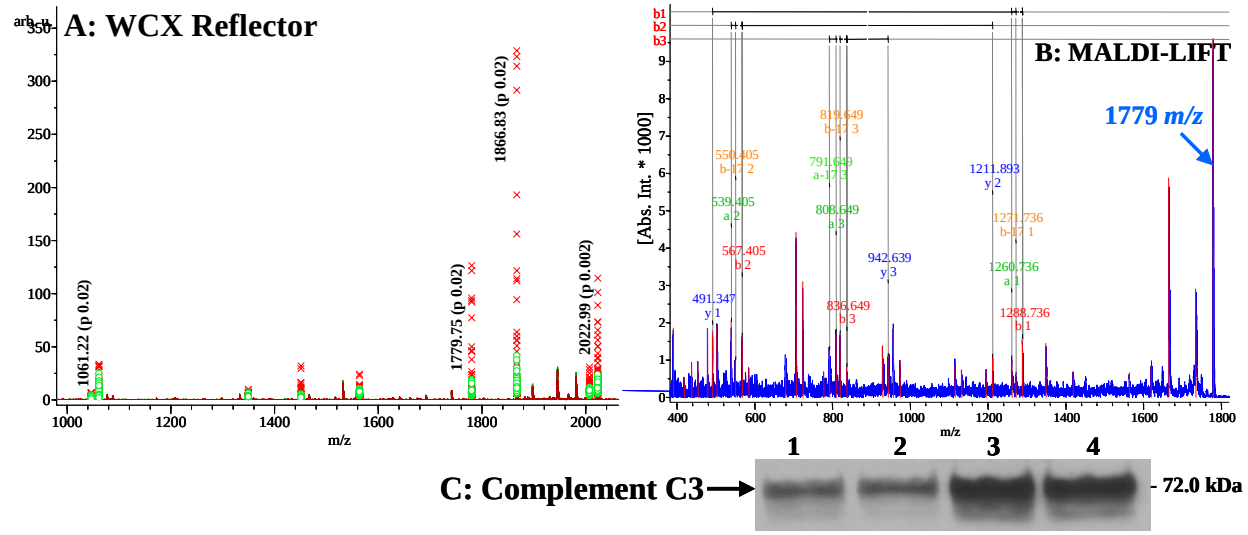


Figure 3.

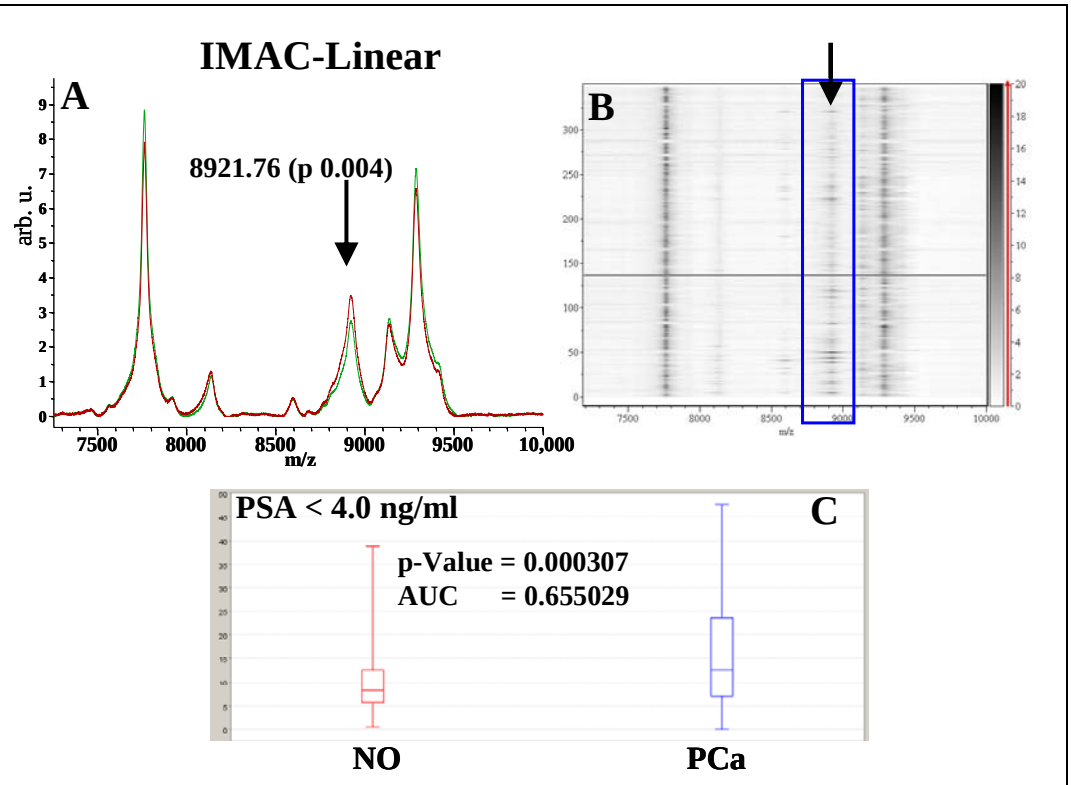


Figure 4.